

Bielecki, Andre; Albers, Sönke; Mantrala, Murali

Working Paper

Salesperson Efficiency Benchmarking Using Sales Response Data: Who is Working Hard and Working Smart?

Arbeitspapiere des Instituts für Betriebswirtschaftslehre, Universität Kiel

Suggested Citation: Bielecki, Andre; Albers, Sönke; Mantrala, Murali (2012) : Salesperson Efficiency Benchmarking Using Sales Response Data: Who is Working Hard and Working Smart?, Arbeitspapiere des Instituts für Betriebswirtschaftslehre, Universität Kiel, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<http://hdl.handle.net/10419/57427>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Salesperson Efficiency Benchmarking Using Sales Response Data: Who is Working Hard and Working Smart?

*André Bielecki*¹

*Sönke Albers*²

*Murali Mantrala*³

¹ Andre Bielecki, Department of Innovation, New Media and Marketing, Kiel University, Westring 425, 24098 Kiel, Germany

² Prof. Dr. Dr. h.c. Sönke Albers, Professor for Marketing and Innovation, Kühne Logistics University, Brooktorkai 20, 20457 Hamburg, Germany

³ Prof. Murali Mantrala, Sam M. Walton Distinguished Professor of Marketing, University of Missouri, 438 Cornell Hall, MO 65211, Columbia, USA

Abstract:

A key to enhancing sales force productivity is finding ways to help existing reps sell more. In this paper, we focus on the process of internal efficiency benchmarking of a firm's sales representatives aimed at identifying strong and weak performers and providing meaningful and actionable directions for improving productivity of relatively inefficient performers. We propose to do this by utilizing measures of two fundamental attributes of a salesperson's controllable work activity as inputs in a DEA (data envelopment analysis) – based procedure: how hard and how smart s/he works. The suggested metrics are derived in an empirical application using archival sales response data from a pharmaceutical company sales force. The application shows that, on average, working smart has larger effects on sales than working hard. In comparison to a conventional DEA benchmarking that simply uses raw sales calls as input measures, the proposed model that uses more 'processing' of the sales response data to derive working smart and hard input measures shows much larger potential for efficiency improvement and offers more meaningful and actionable guidance for improving sales force productivity.

Keywords:

Working Smart, Working Hard, Salesperson Benchmarking, Data Envelopment Analysis, Efficiency Analysis

1. Introduction

In many industries, expenditures on the sales force constitute the single largest marketing budget item for many firms (Mantrala, Albers, Gopalakrishna, and Joseph 2008; Albers, Mantrala, and Sridhar 2010). For example, the pharmaceutical industry in the United States is reported to have spent over \$15 billion on personal selling (detailing) in 2009-2010 (Cegedim-SK&A 2011), and employs well over 75,000 sales representatives (Baldwin 2011). Without doubt, improving sales force productivity is a matter of major concern for the management of these firms.

Traditionally, sales management has viewed increases in sales force size along with judicious changes in salespeople's deployment across markets (e.g., customer groups, products, geographic territories) as crucial steps in enhancing sales force productivity. Accordingly, many models for optimizing sales force size and sales resource allocation have been proposed and implemented over the last few decades (see, e.g., Albers and Mantrala 2008 for a review). However, in tough economic and budgetary times, there is a growing recognition that enhancing sales force productivity is not simply a matter of hiring, firing, and deploying the right number of individuals, but is also critically dependent on *finding ways to help existing reps sell more*. According to some analysts, narrowing the gap between the top 15% or 20% and the rest of the sales force can lead to productivity jumps of over 200% (e.g., Ledingham, Kovac, and Simon 2006). Therefore, it is not surprising that sales managers are devoting increasing attention to individual performance evaluation and *internal benchmarking*, i.e., the process of using internal company data to identify strong and weak performers in the sales force, determine and explain the reasons for the gap between them, and take corrective action to close the gap (e.g., Parsons 2004). In particular, successful selling methods used by top salespeople can be incorporated into company training programs for the development of individual salespeople.

Benchmarking, in general, is a widely-adopted management concept or 'tool' by which a firm seeks to identify and replicate "best practices" to enhance its business performance (Rigby and Bilodeau 2007; Camp 1995; Zairi 1998). There are two broad categories of benchmarking, external and *internal*. While external benchmarking compares the performance of one organization with others in the same industry or across industries, internal benchmarking compares similar operations (individuals, departments, branches etc.) within an organization. Within marketing, selling capability in

particular has been identified as a key target for external benchmarking, e.g., Vorhies and Morgan (2005). In this paper, we focus on internal benchmarking of a firm's sales representatives, i.e., assessing and comparing the *performance efficiency* (in converting their resource inputs into valuable output/s) of the firm's salespeople. This is much easier to implement because the firm can use its own archival data and does not have to rely on the willingness of other comparable firms to exchange data. Current internal sales force efficiency benchmarking procedures in practice tend to be ad hoc rather than systematic, frequently using simple objective output-to-input ratios, e.g., orders per call, expenses per call etc., to assess and compare individual salespeople's performance efficiency (e.g., Johnston and Marshall 2010, 184-213). Such an approach often results in multiple measures of performance efficiency for each rep with mixed results, e.g., a rep may be very efficient relative to his or her peers on one measure while being relatively inefficient on another measure. Subsequently, it is difficult for management to provide consistent guidance to salespeople for improving their performance.

To facilitate comparisons of salespeople's performances and the provision of meaningful, actionable feedback to weaker performers, a salesperson benchmarking technique should provide a single quantitative measure of overall efficiency upon which salespeople can be compared, simultaneously taking into account multiple relevant inputs and outputs (Hershberger, Osmonbekov, and Donthu 2001). Data Envelopment Analysis (DEA, e.g., Charnes, Cooper, and Rhodes 1978) is a technique that produces such an efficiency measure. Specifically, DEA is a linear programming methodology that uses objective data on multiple inputs and outputs as its sole means of producing efficiency scores, e.g., Kamakura, Lenartowicz, and Ratchford (1996), and has already been applied to relative performance assessment of salespeople (e.g., Mahajan 1991; Boles, Donthu, and Lohtia 1995; Pilling, Donthu, and Henson 1999; Hershberger, Osmonbekov, and Donthu 2001). Boles, Donthu, and Lohtia (1995) conclude that their proposed DEA-based approach can prove useful in mentoring and training of the sales force based on the best practices of the most efficient salespeople.

There is little evidence, however, that the application of DEA in sales force evaluation has become widespread. This could be because of inattention to the managerial relevance and meaningfulness of the input and output metrics used in benchmarking – even though it is known that the choice of these metrics is crucial to successful benchmarking analysis and implementation (e.g., Pettijohn et al. 2001). The common

admonition that the metrics used in performance evaluation be controllable by the targets (i.e., sales reps), and impact future outcomes (e.g., Hauser and Katz 1998; Rackham and DeVincentis 1999) is often ignored. For example, Boles, Donthu, and Lohtia (1995) use the following input factors in their benchmarking analysis: the salary of the individual salesperson, the number of sales training sessions, salesperson seniority in months, and the related first line sales manager's management control span. However, none of these input factors are directly influenced by the observed salesperson, and, therefore, these factors are not particularly useful or actionable once the benchmarking exercise is completed. Similarly, Pilling, Donthu, and Hensen (1999) use measures of "territory richness" (total market demand for the industry in question, average sales per account, and growth rate of market demand for that industry) as inputs in their analysis. However, once territories have been designed and assigned, these are more appropriately viewed as *uncontrollable* resources (e.g., Mahajan 1991) from the perspective of the assigned salesperson. As Rackham and DeVincentis (1999, p. 281) put it, if you measure things that salespeople cannot control then your measures will never result in improved sales. Clearly, the sophistication and power of the benchmarking methodology is of little value if the resulting findings on relative efficiencies are ignored due to the lack of meaningful and actionable metrics going into or emerging from the exercise.

Against this backdrop, in this paper, we propose that internal sales force efficiency benchmarking can be made more meaningful and actionable by utilizing measures of two fundamental attributes of a salesperson's controllable work activity as inputs: how *hard* and how *smart* s/he works. These concepts, of course, are not new. Previous research in sales management has suggested working hard and working smart have strong relationships with productivity, e.g., Sujan (1986), Sujan, Weitz, and Sujan (1988), Sujan, Weitz, and Kumar (1994), and Rapp et al. (2006). However, they have not been employed as input factors in sales force efficiency benchmarking, so far.

Working hard is most often conceptualized as the overall amount of effort or time that salespeople devote to trying to achieve sales goals (e.g., Sujan 1986; Sujan, Weitz, and Kumar 1994) (*Effort* itself is the force, energy, or activity put into selling, e.g., Brown and Peterson 1994). On the other hand, *working smart* is typically seen to have two aspects: *sales planning*, e.g., the intelligent deployment of effort across customers, and *adaptive selling* or *communications* to customers, i.e., "*..altering of sales behaviors*

during a customer interaction or across customer interactions based on perceived information about the nature of the selling situation" (Weitz, Sujan, and Sujan 1986, p. 175). More generally, in empirical studies of these two behaviors to date, *working hard* has been measured using archived data, e.g., call records (Rapp et al. 2006). The different aspects of *working smart*, however, have been measured by combining various subjective sub-scales for adaptive selling and planning (e.g., Spiro and Weitz 1990; Sujan, Weitz, and Kumar 1994; Rapp et al. 2006). Until now, measures of *smart* work derived from more objective and readily available archival data have not been developed.

Our proposed approach utilizes measures of hard and smart work derived from historical sales response data as inputs, with observed sales volume as output, in a DEA to identify salespeople who are most efficient (efficiency of 100%) in applying their ‘energy’ and intelligence or ‘smarts’ to produce sales, and can serve as role models (references) for bettering the performance of less efficient salespeople. More specifically, following previous literature, a salesperson’s sales output depends on how hard s/he works, e.g., the *total number of calls* that s/he directs at her/his customers, as well as how smart s/he works. In our conceptualization, the latter is reflected by the optimality of the salesperson’s *allocation* (planning) of calls across sales coverage units (SCUs) and customer accounts or segments comprising her/his geographic territory, as well as *the effectiveness of her/his communications to each customer* (e.g., Rapp et al. 2006). All three metrics can be quantitatively derived from the same panel sales response data that has been previously used in models for sales resource allocation decisions (e.g., Lodish 1988; Skiera and Albers 1998). We propose that salesperson benchmarking constitutes an equally important – but hitherto neglected – use for such panel data. Moreover, we hypothesize that efficiency benchmarking incorporating these input metrics reflecting hard and smart work will lead to more discriminating identification of top performers as well as to more diagnostic and meaningful directions for improving the performance of relatively inefficient salespeople. We derive the suggested input metrics and investigate their benefits for internal efficiency benchmarking in an empirical application involving a pharmaceutical company sales force.

The remainder of this paper is organized as follows. Section 2 describes the panel data, made available to us by a German pharmaceutical company, that we propose to exploit for salesperson benchmarking, and the development of corresponding measures of

hard and *smart* work of individual reps. Section 3 describes the DEA based relative efficiency assessment procedure utilizing the specified input and output measures (e.g., Charnes, Cooper, and Rhodes 1978; Simar 1996; Simar and Wilson 1998, 2000). Section 4 demonstrates the application of the proposed benchmarking approach to the German pharmaceutical firm's sales force. In Section 5 we compare the results of our proposed benchmarking model to those derived from a 'conventional' DEA that uses the same raw output (sales) and input (calls) data. The comparison demonstrates the benefits of the proposed model. The paper concludes with a summary of the main takeaways and directions for future research.

2. Panel sales response data-based metrics for *hard* and *smart* work in field selling

We use longitudinal information about each salesperson's calls on customers in her/his territory and resulting sales to derive the proposed measures of hard and smart work. Below, we describe the data setting, the formulae for the input measures, and the approach for estimating them.

Data Setting

We use panel data from a medium-sized German pharmaceutical company whose products are promoted by a sales force. The firm has an extensive database, which records all sales calls to four pre-specified physician segments (VIP, A, B and C), which were formed according to their prescription potentials. The firm also obtains data on sales from IMS Health. However, to comply with German privacy legislation, market information services like IMS are barred from providing individual physician-level sales data, and can only provide prescription data at a higher level of aggregation, namely sales coverage unit or SCU-level data. Each SCU represents an aggregation of at least six physicians, so that no data can be traced back to individual physicians. Our collaborating firm has access to IMS Health sales data and pharmacy coverage for 1860 sales coverage units. Further, the company has 55 sales territories in Germany, each aligned to a single salesperson, with 34 SCUs per territory on average. The data contain multiple observations for each SCU over 43 months (from years 2001 to 2004). The multiple data points allow individualized (salesperson-specific) statistical response estimation. Subsequently, we use the first 36 months of the panel (2001 to

2003) for estimation purposes, and hold out the remaining seven months (2004) for model validation.

Working smart metrics

1. *Call effectiveness*

Our metric for call effectiveness is *the effect of selling per salesperson* as obtained from econometrically estimated semi-log response function parameters for each sales territory (see Appendix Section A). This represents an effectiveness measure because it tells us how much sales can be increased if selling effort is increased by one unit. It is also comparable across units. We estimate these idiosyncratic sales response functions using a hierarchical Bayesian linear model as described by Rossi, Allenby, and McCulloch (2005). The salesperson-specific response coefficients reflect the individual salesperson's *call effectiveness*. More specifically, as the physicians in the segments s ($s=1, \dots, S$) covered by each salesperson g ($g=1, \dots, G$) have different sales potentials, we use the sum of the salesperson-segment response coefficient estimates *Coefficient* _{g,s} , each weighted by the respective segment sales potential, *SegPot* _{g,s} , as the measure of her/his overall call effectiveness, *CE* _{g} . That is,

$$CE_g = \sum_s^S SegPot_{g,s} \cdot Coefficient_{g,s} \quad (1)$$

The segment potential *SegPot* _{g,s} is calculated by the number of assigned sales accounts in segment s , in the observed territory g , multiplied by a segment specific potential factor for segment s which is given by the data providing pharmaceutical firm.⁴

2. *Effort allocation quality*

Using each salesperson's estimated sales response model, we can determine the *optimal allocation* of a salesperson's details across customer segments and SCUs and

⁴ As an example, in a salesperson's territory g the segment VIP contains 17 physicians (segment A: 70, segment B: 46, and segment C: 13 physicians). The data providing firm multiplies segment VIP with a multiplier of 2.356 reflecting its selling potential (multiplier for segment A: 1.292, segment B: 0.773, and segment C: 0.305). Therefore, the segment potential for segment VIP in the example territory is $SegPot_{g,VIP} = 17 \cdot 2.356 = 40.05$ ($SegPot_{g,A}: 70 \cdot 1.292=90.44$, $SegPot_{g,B}: 46 \cdot 0.773=35.56$, $SegPot_{g,C}: 13 \cdot 0.305=3.97$).

compare the resulting sales volume OA_g with the sales volume AA_g derived from the *actual allocation* (provided in the panel data). Then, for each salesperson g , the *effort allocation quality*, AQ_g , reflecting the ‘smartness’ of the rep’s actual allocation, is given by:

$$AQ_g = OA_g - AA_g \quad (2)$$

Working hard metric

How hard a salesperson works is given by $Effort_g$, the average number of calls per month s/he makes over the performance assessment horizon of T months. More specifically:

$$Effort_g = \sum_r^R \sum_s^S \frac{1}{T} \sum_t^T (Calls_{r,s,t}) \quad (3)$$

where, $Calls_{r,s,t}$ denotes the calls in month t ($t=1, \dots, T$), to customer segment s in SCU r ($r=1, \dots, R$).

Sales territory potentials

In addition to the above three salesperson inputs, we include the reps’ territory potentials as an uncontrollable input factor.⁵ The g -th salesperson’s territory sales potential, $TerrPot_g$, is computed as follows:

$$TerrPot_g = \sum_s^S SegPot_{g,s} \quad (4)$$

$SegPot_{g,s}$ is the physician segment s and territory g specific sales potential from equation (1).

⁵ Uncontrollable means that DEA accounts for territory specific differences in the sales potential, but for the calculation of efficiency scores, and for the improvement of individual reps, the potentials are assumed to be unchangeable for the salesperson, i.e., the uncontrollable potential factor does not need to be changed by the salesperson to achieve full efficiency. (In contrast controllable factors like the *call effectiveness*, the *call allocation quality* and the *effort* may need to be changed by the salesperson to achieve efficiency).

3. Salesperson benchmarking using Data Envelopment Analysis

As explained in previous works (e.g., Dutta, Kamakura, and Ratchford 2004), DEA determines the relative efficiency with which each unit (salesperson in our case) under evaluation converts or transforms multiple inputs into single or multiple outputs of interest, where efficiency is the ratio of weighted output/s to weighted input/s. More specifically, DEA determines the unit-specific input and output weights that maximize each unit's efficiency subject to all units' efficiencies using these same weights being less than or equal to one (or 100%). Thus, DEA inherently compares the performances of all units by deriving the unit-specific input and output factor weights that are the basis for its relative efficiency score. Those units whose optimal weights result in an efficiency score of 100% determine the 'efficiency frontier' for the collection of units under evaluation. Units falling on the efficiency frontier are those who compared to other units provide no evidence of inefficiency in their conversion of inputs to outputs. A subset of these efficient reference units serves as 'benchmarks' (or best practice set) for each inefficient unit (Charnes, Cooper, and Rhodes 1978). By combining the 'technologies' of their benchmark units, each inefficient unit should be able to produce her/his current sales using only a fraction of her/his current inputs equal to her/his relative efficiency score, $efficiency_g$. Alternatively, a sales volume can be computed which has to be reached in order to be comparable with benchmarking units using the same weighted combination of inputs. The target *sales objective*, SO_g , that must be achieved by each salesperson to attain 100% efficiency with her/his current level of inputs is:

$$SO_g = y_g \cdot \frac{1}{efficiency_g} \quad (5)$$

where y_g is her/his current sales level.

Lastly, our analysis supplements the output-oriented basic DEA model (Charnes, Cooper, and Rhodes 1978) with bootstrapping for the estimated efficiency scores as described by Simar (1996), and Simar and Wilson (1998, 2000). This 'stochastic DEA' or SDEA applies bootstrapping to enhance the linear programming DEA results with a statistical underpinning and calculates the bias in the efficiency of each rep. The bias correction gives additional credibility to the efficiency scores and supports the

stability of the efficiency benchmarking. Details on SDEA and on the bootstrap computations can be found in Appendix Section D.

The next section describes the steps and results of a SDEA application for a pharmaceutical sales force based on working hard and working smart input metrics.

4. Application

Individualized sales response estimation and achieved quality of call response

The individualized sales response estimation is performed using the hierarchical Bayesian linear model proposed by Rossi, Allenby, and McCulloch (2005). The model is estimated using the *bayesm* package in R (Rossi 2011).

For each rep we estimate an individual set of response coefficients. A rep's sales volume is used as dependent variable. Therefore, the estimated set of coefficients describes how much the sales volume can be increased in the observed rep's territory (which is the unit under consideration) if the rep puts in one more unit of selling effort. As independent variables, we include stock variables of a rep's detailing effort across four customer potential segments (VIP, A, B, C). The territory pharmacy (under- or over-) coverage is used as an additional variable reflecting the specialties of sales closure in the pharmaceutical industry (sales 'take place' at pharmacies, not at the physicians). A time trend is captured by a linear trend variable⁶. More details on the sales response model specification and its estimation can be found in the Appendix Section A.

Response model estimation results

Utilizing the first 36 months of data (we hold out 7 months' data from 2004 for model validation), seven sales response model coefficients (per salesperson) are estimated on a disaggregated level, i.e., an intercept, four physician segment-specific sales-to-call coefficients, a pharmacy coverage response coefficient, and a time trend coefficient, for each individual salesperson, Table 1 displays point estimates and t-values for the

⁶ A linear trend has proven to be the best fitting specification.

estimated salesperson-segment sales response coefficients for five sample sales reps⁷, the means of estimated model coefficients across all 55 sales reps, and the proportions of these coefficients that were not significant at the 1%, 5%, and 10% levels. (A non-disclosure agreement with the company prevents us from displaying all 55 reps' estimated coefficients.)

Table 1: Estimated coefficients and t-values

Five example reps	Intercept		Sales response coefficients for physician segments VIP to C								Sales response coefficient for pharmacy coverage		Sales response coefficient for time trend	
			Segment VIP		Segment A		Segment B		Segment C					
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
1	5205***	4.86	32.74***	15.01	54.82***	28.96	64.33***	9.00	65.62*	1.78	56.47***	13.56	12.04***	14.30
2	4093***	4.69	46.47***	20.52	55.70***	23.23	63.15***	12.12	49.71**	2.52	24.85***	12.27	10.08***	12.78
3	7424***	6.50	30.07***	11.75	68.12***	25.22	21.74***	3.30	126.03***	6.16	38.98***	12.15	11.00***	10.42
4	6737***	7.22	38.35***	33.18	76.71***	41.71	47.24***	9.58	211.38***	8.77	82.74***	12.48	9.43***	11.89
5	12956***	14.47	40.36***	23.08	66.33***	32.70	27.80***	7.42	-3.72	-0.31	51.64***	6.91	4.50***	8.97
...	
Mean (for all 55 reps)	3636		44.89		56.32		54.66		88.92		43.31		11.55	
% of rep coeff. not sig. at 1% level	30.91%		0.00%		0.00%		12.73%		29.09%		1.82%		0.00%	
% of rep coeff. not sig. at 5% level	27.27%		0.00%		0.00%		5.45%		23.64%		1.82%		0.00%	
% of rep coeff. not sig. at 10% level	20.00%		0.00%		0.00%		3.64%		18.18%		1.82%		0.00%	

*significant at 10% level, **significant at 5% level, ***significant at 1% level, one-sided tests, adjusted R² = 97.9%

Overall, a high proportion of the model estimation results across the 55 reps is significant with plausible signs and magnitudes. The significance results indicate a high stability for the estimated weights and the goodness of fit of the overall model to the data is high. Moreover the estimated model predictions of sales in each of the seven hold-out months are good. Specifically, comparing each rep's predicted and actual sales results for the seven months, the average Mean Absolute Percentage Error (MAPE) over all 55 sales representatives is 18.15%. Formal details on the calculations can be found in the Appendix Section B. The prediction errors range from 5.31% for Rep 4 to

⁷ Here and in the following tables the sales rep case ID numbers (territories) have been arbitrarily re-numbered so that results cannot be traced back to individuals. Also, we only report results for a few reps due to nondisclosure restrictions.

about 48% for Rep 27; the standard deviation is 11.01%. Overall, the MAPE measures suggest the response model estimation results have reasonable predictive validity.

Assessments of call effectiveness based on estimated response coefficients

Using the estimated response coefficients, the working smart measure of call effectiveness was calculated for each salesperson according to Equation (1). Table 2 displays the minimum and maximum, the mean and the standard deviation of this metric. The call effectiveness ranges from 5,506 to 18,516, the sales force average is 11,491, and the standard deviation 2,856.

Table 2 also displays normalized values (scaled from 0 to 1) for the *call effectiveness* and the two remaining input metrics *effort allocation quality* and *effort*. In DEA benchmarking the normalization is useful for comparing the input factor weights across factors. Without normalization of input, the factors have very different dimensions, e.g., call effectiveness is measured in ‘thousands’, but effort is measured in ‘tenth’. Thus, normalization makes the resulting input factor weights comparable across factors. By the interpretation of (normalized) input factor specific median values of DEA factor weights across salespersons, we can say that the factor with the highest median weight has, on average across salespersons, the largest effect on the generation of sales.

Table 2: Salespersons' call effectiveness, allocation quality, and effort

	Call effectiveness	Allocation quality <i>Sales volume difference (optimized minus actual sales volume)</i>	Effort
<i>Non normalized values</i>			
Min.	5,506	664	27
Max.	18,516	12,839	125
Mean	11,491	3,674	91
SD	2,856	2,638	14
<i>Normalized values (0 to 1)</i>			
Min.	0.00	0.00	0.00
Max.	1.00	1.00	1.00
Mean	0.64	0.25	0.48
SD	0.26	0.22	0.16

Assessments of salesperson effort allocation quality

The individual effort allocation quality is calculated as per Equation (2) in the previous section. We use the estimates of the rep- and segment-level sales response coefficients to calculate the sales volume achievable by the optimized (sales-maximizing) allocation of calls across SCUs and physician segments (VIP, A, B, and C), e.g., Skiera and Albers (1998). We optimize the sales volume, rather than e.g., the territory profit, because there were no margin differences across territories for the underlying problem and salespersons are usually assessed according to the sales territory volume. We use *Excel Solver* for this optimization analysis (formal details can be found in the Appendix Section C). Reps with a smaller gap between their actual and optimized sales volumes allocate their calls in a smarter way. Table 2 displays the minimum and maximum, the mean and the standard deviation on this metric (and the normalized values). We note that the sales force average effort allocation quality (the average monthly gap) is €3,674, and the standard deviation is €2,638. If all salespeople successfully achieve their optimal effort allocations, the overall monthly sales volume will increase by over €202,000 (12.29%), i.e., an improvement of nearly €2,425,000 per year.

Salesperson level of effort

The individual salesperson effort is calculated according to Equation (3). Table 2 displays the minimum and maximum, the mean and the standard deviation on this metric (and the normalized values). The territory-specific effort levels range from 125 calls per month to about 27 calls per month. The sales force average is about 91 and the standard deviation is 14.12.

Stochastic Data Envelopment Analysis results

Performing the SDEA with sales reps' average monthly sales volume as output, and their estimated call effectiveness, effort allocation quality, effort level, and uncontrollable territory sales potential as inputs, we find there are *ten* (100%) efficient salespersons (out of 54 reps, one who was a clear outlier was dropped from this analysis⁸). That is, 44 reps were deemed inefficient to varying degrees, implying there is much scope for improvement in this exemplary sales force. The Tables 3 and 4 show the DEA results.

The ten top performing reps are 1, 14, 17, 21, 28, 40, 44, 45, 53, and 55. Column 3 of Table 3 displays the results for the five most efficient and inefficient reps.⁹ The latter's relative efficiency scores are rather low, ranging from 38.5% to 23.3%. Column 4 of Table 3 displays the *efficiency bias*, derived from (300) bootstrap replications. Thus, the mean efficiency score for all 54 evaluated reps is $65.7\% \pm 4.63\%$. Next, Columns 6 to 9 of Table 3 show the four input factors' (call effectiveness, effort allocation quality, effort, and territory potential) normalized values on a range from 0 to 1 (remember that we normalized the input factors for better comparability of factor weights across factors). The DEA-determined input factor weights, reflecting the relative contribution of the factor to the salesperson's generation of sales output, are displayed in Table 3, Columns 10 to 13. We see that, on average, effort allocation quality (median weight of 3.17) has the largest effect on sales with territory potential having the second largest effect (median weight 0.84), followed by call effectiveness (median weight 0.76) and

⁸ One salesperson (territory) was sorted out because the available data suggest that the territory was not occupied by a salesperson for a large share of the available month. E.g. the sales volume and the number of calls are extremely small and may bias the DEA results.

⁹ Remember that the numbers of the territories have been arbitrarily renumbered so that results cannot be traced back to individuals. In addition, we only report results for some territories due to the same concern.

the effort level (median weight 0.34). Thus, on average, working smart (effort allocation quality and call effectiveness) has a larger effect on sales in comparison to working hard (effort). This is a remarkable result. The outstanding importance of salespersons' smartness for generating sales should be of major concern for sales force management. No previous sales force related analysis has pointed out this large importance of salesperson smartness, also in direct comparison to working hard (which, actually, seems to be less important in our application).

Table 3: Results of the proposed DEA model: Efficiency and importance of factors

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Rank	Salesperson	Efficiency score	Efficiency bias	Sales volume (€)	Input: Call effectiveness	Input: Effort allocation quality	Input: Effort	Input: Territory potential	Weight: Call effectiveness	Weight: Effort allocation quality	Weight: Effort	Weight: territory potential
1	14	100.0%	0.0%	36,017	0.15	0.27	0.01	0.34	0.00	0.00	5.37	0.00
1	28	100.0%	0.0%	31,634	0.37	0.08	0.33	0.39	1.23	3.11	0.00	0.35
1	40	100.0%	0.0%	30,819	0.45	0.20	0.31	0.04	1.70	0.00	0.38	2.71
1	45	100.0%	0.0%	34,882	0.73	0.03	0.41	0.35	0.00	14.93	0.00	1.08
1	55	100.0%	0.0%	31,297	0.49	0.14	0.46	0.01	0.00	0.00	0.00	27.80
...												
50	32	38.5%	2.1%	20,904	0.71	0.16	0.58	0.52	0.97	4.19	0.95	1.32
51	35	38.2%	3.5%	24,203	0.93	0.13	0.85	0.69	1.21	5.02	0.00	1.20
52	4	37.6%	8.7%	25,493	0.34	0.47	0.46	0.67	2.53	3.84	0.00	0.00
53	8	34.1%	9.1%	27,006	0.61	0.57	0.38	0.51	2.07	0.10	0.28	2.96
54	27	23.3%	1.3%	24,113	0.84	0.55	1.00	0.98	1.73	3.25	0.00	1.07
Mean		65.7%	4.6%	30,081	0.64	0.25	0.48	0.54	0.76*	3.17*	0.34*	0.84*

*Median values because of some substantial outliers (e.g., weight of >100 for effort allocation quality)

Moreover, our application shows that no new data is required for the analysis; rather archival sales response data can be used for analyzing working smart and hard. The requirements that input factors should be controllable for the salesperson and directly affect future outcomes (sales) are met as well, because 'planning call allocation' and 'training call effectiveness' (working smart) as well as 'doing calls' (working hard) are major parts of a salesperson's regular activities.

It is also interesting to look at the factor weights of some individual salespersons. On average, working smart (effort allocation quality and call effectiveness) has the largest effect on sales (weights 3.17 and 0.76), but individually some salespersons have other priorities. As an example efficient Rep 14 has very large weight on effort (5.37) (working hard) and nearly no weight on call effectiveness and effort allocation quality (working smart). A contrary example is efficient Rep 45 who has very strong weight on the effort allocation quality (14.93) (working smart) and nearly no weight on the other factors (there is only some additional weight of 1.08 on the territory potential).

Again, for the complete sales force working smart, i.e., effort allocation quality and call effectiveness, have larger effects on sales in comparison to working hard, i.e. calling effort. This is remarkable because up to now only little work has been put into using quantitative data for analyzing salespersons' smartness. It has already been noted by Sujan (1986), Sujan, Weitz, and Kumar (1994) and Rapp et al. (2006) that working smart and hard have strong relationships with productivity. Thus, managers should try to put more work on the optimization of salespersons smartness, e.g., by training or motivation.

While Table 3 displays the factors and the DEA factor weighting, Table 4 describes the benchmarking relations, i.e., the orientation of inefficient reps at their respective efficient 'role models'. Again, Columns 1 to 4 show the ranks, the salespersons, the efficiency scores and the sales volumes. Column 5 displays the sales objectives for each individual salesperson according to Equation (5). If a salesperson achieves her/his sales objective, s/he becomes fully efficient (100%) in the meaning of her/his composite benchmark. Columns 6 and 7 show the potential for improvement for the individual inefficient salesperson in absolute numbers (€) and in percent. Of course, such (very large) improvements are not likely to be fully realized in practice, but indicate the potential gains for each individual salesperson. Again, the last line shows the mean values for the complete sales force.

Comparing the results in Table 4 Columns 4 and 5 we see that the actual average monthly sales volume is €30,081 compared to the average monthly achievable *sales objective* of €50,335 if all salespeople became fully (100%) efficient. Aggregated to the complete sales force and one year, this represents an overall potential sales volume improvement of more than €13 million. Therefore, the potential gain from inefficient reps' emulation of their respective benchmarks is very high.

Table 4: Results of the proposed DEA model: Benchmarking

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rank	Salesperson	Efficiency score	Sales volume (€)	Sales objective (€)	Potential for improvement (€)	Potential for improvement %	Benchmarks	No. of influenced salespersons
1	14	100.0%	36,017	36,017	0	0%		39
1	28	100.0%	31,634	31,634	0	0%		31
1	40	100.0%	30,819	30,819	0	0%		2
1	45	100.0%	34,882	34,882	0	0%		29
1	55	100.0%	31,297	31,297	0	0%		21
...								
50	32	38.5%	20,904	54,227	33,323	159%	14, 28, 45, 55	
51	35	38.2%	24,203	63,405	39,202	162%	28, 45, 55	
52	4	37.6%	25,493	67,814	42,321	166%	14, 28	
53	8	34.1%	27,006	79,241	52,235	193%	14, 21, 40, 55	
54	27	23.3%	24,113	103,353	79,240	329%	14, 28, 55	
Mean		65.7%	30,081	50,335	20,254	72%		

But, how can these improvements be achieved? In Table 4 Column 8 indicates which reps may be used as benchmarks for each observed inefficient rep. We see that each inefficient rep can have several benchmarks, e.g., rep 32 has four, namely, reps 14, 28, 45, and 55. Conversely, Column 9 of Table 3 indicates the number of inefficient reps for which each efficient rep serves as a reference. For example, rep 14 is a benchmark for 39 other salespersons.

For a better understanding we discuss the benchmarking of inefficient rep 4 on rank 52 as an example. Inefficient Rep 4 has two benchmarks, the efficient rep 14 and the efficient rep 28. In terms of DEA, benchmarks achieve higher efficiency using the factor weights of the observed (inefficient) rep and their own factor levels (Charnes, Cooper, and Rhodes 1978; Avkiran 1999). In an immediate comparison of input and output factors between inefficient rep 4 (Table 3: call effectiveness: 0.34, effort allocation quality: 0.47; effort: 0.46, territory potential: 0.67) and her/his benchmarks reps 14 and 28, we see that both benchmarks use less input (less call effectiveness: rep 14: 0.15, rep 28: 0.37, less effort allocation quality: rep 14: 0.27, rep 28: 0.08, less effort: rep 14: 0.01, rep 28: 0.33, and have less territory potential: rep 14: 0.34, rep 28: 0.39) and

nonetheless achieve higher output (sales volume). Therefore, using her/his (larger) input rep 4 should be able to achieve higher sales (higher than her/his own achieved sales and also higher than the sales of benchmark reps 14 and 28). Otherwise, to be efficient rep 4 should use (need) less input for her/his actual sales volume.

The example shows, that in comparison to efficient reps 14 and 28, rep 4 shows inefficient input factor utilization for generating sales. For her/his 'rather high' level of *hard work* (effort) and *smart work* (call effectiveness, call allocation quality) s/he achieves 'rather small' sales.

Next, we should ask how inefficient rep 4 could improve her/his performance? For the generation of sales rep 4 primarily uses *smart work*, i.e., the factors *call effectiveness* (weight: 2.53) and *effort allocation quality* (weight: 3.84). Therefore, for the improvement of her/his efficiency rep 4 may try to improve her/his call allocation or attend further sales training to improve her/his call effectiveness. Rep 4 has no weight on the factor *effort*. In combination with the more or less "average" input factor effort (rep 4: 0.46; salesforce average: 0.48), which does not indicate a strong 'lack of effort' for rep 4, the performance of rep 4 in working hard (effort) does not seem to be a large problem. Thus, the primary task for rep 4 should be the improvement of call effectiveness and call allocation, i.e. her/his smartness in calling on customers.

Rep 4 was just one example, but similar benchmarking can be applied for the remaining inefficient reps. It is important to note, that the clear and actionable advice '*what to improve*' (smart and/or hard work?) and '*what is the potential for improvement*', result from processing of the very basic archival sales response data (calls) to more sophisticated and meaningful working smart and hard input metrics, which reflect the salespersons' regular activities, which are controllable for the salespersons, and which directly affect the sales volume. The meaningfulness and managerial relevance of the input metrics used in benchmarking is crucial for the analysis and a successful implementation (Pettijohn et al. 2001).

A relevant question at this stage is how does our proposed use of working hard and working smart measures derived from sales response data for benchmarking compare with a 'conventional DEA' that would have used simply the raw number of calls as the controllable input factor? The next section examines this question.

5. Comparison of the proposed model to conventional DEA benchmarking

In this section, we compare the results of the proposed benchmarking specification to results based on a conventional DEA, which describes the more or less ‘naive’ way of using the available sales response data for salesperson efficiency benchmarking. The key differences between the proposed and the conventional model lie in the methods *how the available sales response data are processed to input factors, i.e., whether they are processed to meaningful and actionable working smart and hard metrics (proposed model), or whether the model uses raw sales calls as input (conventional model)*. The differences in the results of both models describe the additional value of processing the data to analyze salespersons’ working smart and hard, instead of pure sales calls. Therefore, most of the model specification is the same for both models. The same archival sales response data, the same stochastic DEA model, and the same output factor, i.e. salespersons’ sales volume is used. The uncontrollable input factor territory potential (Equation 4) is also the same for both models. The key differences lie in the controllable input factors, i.e., if the sales response data is processed for salesperson benchmarking in working smart and hard metrics (proposed model) or if the model uses pure sales calls from four customer segments (VIP, A, B C) (conventional model). Table 5 summarizes the key differences.

Therefore, in the conventional model we have the salespersons’ sales volume as output, and five input factors which are the sales calls per salesperson g for segments s (VIP, A, B, and C), and the uncontrollable territory potential. The comparison of the results of both models shows the additional benefits of analyzing whether/how a salesperson works *smart* and/or *hard*. Details on the specification of the input factors for the conventional model can be found in the Appendix Section E.

Table 5: Key differences between the proposed and the conventional benchmarking model

	Proposed benchmarking model	Conventional DEA benchmarking
Data:	Archival sales response data	Archival sales response data
Benchmarking model:	SDEA	SDEA
Output factor:	Salespersons' territory sales	Salespersons' territory sales
Controllable inputs:	<i>Call effectiveness</i> <i>Effort allocation quality</i> <i>Effort level</i>	<i>Calls to customer segment VIP</i> <i>Calls to customer segment A</i> <i>Calls to customer segment B</i> <i>Calls to customer segment C</i>
Uncontrollable input:	Territory Potential	Territory Potential
Annotations:	The proposed model processes the archival sales response data into two fundamental attributes of a salesperson's controllable work activity: how <i>smart</i> and how <i>hard</i> s/he works. The processed factors <i>call effectiveness</i> , <i>effort allocation quality</i> , and <i>effort</i> are used as proxies for salespersons' <i>working smart</i> and <i>working hard</i> activity.	The conventional model uses the raw, (unprocessed) sales response data, i.e., sales calls according to four customer potential segments (VIP, A, B, C). The unprocessed factors provide no information on salespersons' <i>smart</i> and <i>hard</i> work, but rather on the distribution of calls to the four customer segments (VIP, A, B, C).

Performing a SDEA similar to that in the last section, we find the average efficiency of the 54¹⁰ salespeople is 85.70%. However, the conventional DEA provides much more favorable assessments of the efficiency of this sales force than indicated by the proposed model (which indicates an average efficiency of 65.73%). While this may be comforting, it is not very diagnostic from the viewpoint of performance improvement. In fact, there are 15 out of 54 efficient reps from this analysis compared to only 10 out of 54 identified by the proposed model. The 15 efficient reps according to the conventional model are Reps 16, 18, 19, 21, 22, 24, 26, 31, 35, 37, 41, 49, 50, 51, and 53. As only inefficient reps can be improved, this conventional DEA indicates much lower potential for sales force performance improvement than the analysis in the previous section. The Tables 6 and 7 summarize the DEA results similar to Tables 3 and 4.

¹⁰ Like in the proposed model, one outlier salesperson was sorted out, so that the total number of salespersons reduces from 55 to 54.

Table 6: Results for a conventional DEA model: Efficiency and importance of factors

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Rank	Salesperson	Efficiency score	Efficiency bias	Sales volume (€)	Input: Calls VIP	Input: Calls A	Input: Calls B	Input: Calls C	Input: Territory potential	Weight: Calls VIP	Weight: Calls A	Weight: Calls B	Weight: Calls C	Weight: Territory potential
1	16	100%	0.00%	30,930	196	513	48	24	201	0.33	0.45	0.15	0.07	0.00
1	18	100%	0.00%	30,984	231	587	27	16	206	0.00	0.00	0.82	0.18	0.00
1	19	100%	0.00%	33,735	160	431	142	34	189	0.14	0.59	0.08	0.19	0.00
1	21	100%	0.00%	37,566	313	403	68	58	221	0.00	0.71	0.29	0.00	0.00
1	22	100%	0.00%	28,750	103	572	24	78	199	0.72	0.00	0.28	0.00	0.00
...														
50	6	67.49%	4.41%	27,083	207	486	225	93	187	0.00	0.09	0.00	0.20	0.71
51	43	67.47%	4.78%	25,799	119	480	169	70	214	0.12	0.66	0.07	0.16	0.00
52	39	65.31%	3.20%	27,915	202	550	160	88	215	0.00	0.03	0.13	0.17	0.67
53	5	66.82%	7.01%	28,667	129	446	236	148	211	0.25	0.75	0.00	0.00	0.00
54	30	70.42%	12.11%	26,938	90	505	174	102	198	0.47	0.53	0.00	0.00	0.00
Mean		85.70%	3.37%	30,081	207	486	225	93	187	0.13	0.37	0.12	0.14	0.24

Column 3 of Table 6 displays the results for the five most efficient and least efficient reps¹¹ and Column 4 displays the efficiency bias, derived from (300) bootstrap replications. Next, Columns 5 to 10 show the output (sales volume) and input factors (sales calls in segment VIP, A, B, C, and territory potential). The DEA-determined input factor weights are displayed in Columns 11 to 15. We see that on average the sales calls in physician segment A have the largest effect on sales (the average weight across all salespersons is 0.37).¹² The average weights for calls in the remaining segments (VIP, B, C) have very similar magnitudes (VIP: 0.13, B: 0.12, C: 0.14). Thus, on average the conventional model shows that sales calls in physician segment A have the largest effect on sales.

¹¹ Remember that the numbers of the territories have been arbitrarily renumbered so that results cannot be traced back to individuals. In addition, we only report results for some territories due to the same concern.

¹² In the conventional model it is not necessary to normalize the input factors for comparable weights, because the sales calls in customer segments VIP, A, B, and C already have the same dimensions.

This is an interesting result, but only helps with the decision regarding which segments should be called for the largest effects on sales (Segment A). But, the activities of a salesperson are more than performing calls in customer segments. It is also planning the calls, preparing the calls, training, motivation and other ‘smart’ activities. Thus, the *conventional model provides no guidance with respect to which salesperson works hard and/or smart and how to improve along these two more diagnostic directions. This is only possible by processing the sales response data into the working smart and working hard metrics, utilized in the previous sections, i.e., with the model proposed in Section 4.*

Table 7 describes the benchmarking relations, i.e., directions for efficiency improvement and orientation for inefficient reps at efficient ‘role model’ reps. Again, the first Columns 1 to 4 show the ranks, the salespersons, the efficiency scores, and the actual sales volume. Column 5 displays the sales objectives for each individual salesperson (also calculated according to Equation (5) like in the proposed model). Columns 6 and 7 show the potential for improvement for each individual (inefficient) salesperson in absolute numbers (€) and in percent.

By comparing the results in Table 7 Columns 4 and 5 we see that the average actual monthly sales volume is €30,081 and the average achievable monthly sales objective is €37,064, which is realized if all reps would become fully (100%) efficient. This represents an overall potential for sales volume improvement of about €4.5 million per year for the complete sales force. Nevertheless, this is a rather small improvement, compared to the improvements achievable by the proposed model, offering potential gains of more than €13 million per year. Remember, between the proposed and the conventional model, we only changed the specification of the controllable input factors (working smart/hard vs. sales calls in customers segments VIP, A, B, C). That is, the original sales response data, the DEA specification, and the specification of uncontrollable input factors remain exactly the same. *Thus, the larger potential gains achievable by the proposed model are only based on processing of the original sales response data into more meaningful working smart and hard input factors.*

Table 7: Results for a conventional DEA model: Benchmarking

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Rank	Salesperson	Efficiency score	Sales volume (€)	Sales objective (€)	Potential for improvement (€)	Potential for improvement (%)		No. of influenced salespersons
1	16	100%	30,930	30,930	0	0.00%		7
1	18	100%	30,984	30,984	0	0.00%		12
1	19	100%	33,735	33,735	0	0.00%		17
1	21	100%	37,566	37,566	0	0.00%		18
1	22	100%	28,750	28,750	0	0.00%		5
...								
50	6	67.49%	27,083	42,935	15,852	58.53%	24, 26, 31, 53	
51	43	67.47%	25,799	41,153	15,354	59.52%	19, 26, 41, 51	
52	39	65.31%	27,915	44,937	17,022	60.98%	18, 21, 26, 53	
53	5	66.82%	28,667	47,923	19,255	67.17%	26, 35	
54	30	70.42%	26,938	46,190	19,252	71.47%	26, 51	
Mean		85.70%	30,081	37,064	6,983	24.85%		

In Table 7 Column 8 describes which salespersons may be used as benchmarks for each individual inefficient rep. Again we see that each inefficient salesperson can have more than one efficient benchmark, e.g., inefficient rep. 6 on Rank 50 has four benchmarks, namely the reps 24, 26, 31, and 53. For efficient reps Column 9 indicates the number of inefficient reps for which each efficient rep serves as a reference. The benchmarking procedure takes place the same way described in the discussion of the proposed model in Section 4. Thus, we do not describe the benchmarking details again.

A relevant question at this stage is: *How does the proposed model perform (using working smart and hard metrics) in comparison to the 'naive' approach using raw sales calls as input factors, namely the conventional model.* Table 8 summarizes and compares the key results from both models.

In the conventional model the yearly potential for improvement per salesperson is €83,796. Using working smart and hard input factors as in the proposed model, the yearly potential for improvement per rep increases to €243,048. This makes a very

significant difference of more than €150,000 per year and per rep on the part of the proposed model.

For the complete sales force, the conventional model offers a potential sales improvement of €4.5 million per year (or about €377,083 per month) compared to a total annual potential sales improvement of more than €13 million per year under the proposed analysis.

Therefore, processing the sales response data to meaningful and actionable working smart and hard input metrics (like in the proposed model) means additional potential for sales volume improvements of more than €8.5 million per year for the sales force, compared to the conventional benchmarking specification (€4.5 million vs. €13 million).

Table 8: Comparison of main results of the proposed and conventional benchmarking model

		Proposed model (“working hard & smart”)	Conventional model (“pure calls”)	Difference
Number of efficient salespersons		10	15	5
Share of efficient salespersons [in %]		18.52%	27.78%	9.26 points (50 %)
Average Efficiency Score (corrected)		65.73%	85.70%	19.97 points (30.38%)
Potential for improvement				
Per salesperson	Percentage points (%)	72.40%	24.85%	47.55 points (191.34%)
	€ per month	€20,254	€6,983	€13,271
	€ per year	€243,048	€83,796	€159,252
Total sales force	€ per month	€1,093,690	€377,083	€716,607
	€ per year	€13,124,283	€4,524,996	€8,599,287

A first indication for the larger potential for improvement is also observable in the models’ average efficiency and the share of efficient salespersons. In DEA, only inefficient reps can be improved (Charnes, Cooper, and Rhodes 1978). In the proposed

model, the smaller share of efficient salespersons (proposed model: 18.52%, conventional model: 27.78%) and the smaller average efficiency (proposed model: 65.73%, conventional model: 87.70%) shows that the proposed model offers a larger potential for improvement, because there is simply a larger number of salespersons (and inefficiency) to improve.

6. Summary of contributions, limitations and future research

DEA-based salesperson efficiency benchmarking has been proposed by several previous papers in the sales management literature (e.g., Mahajan 1991; Boles, Donthu, and Lohtia 1995; Pilling, Donthu, and Henson 1999; Hershberger, Osmonbekov, and Donthu 2001). However, this approach to sales performance evaluation does not appear to be common or widespread in sales management practice. We believe that much of this resistance to utilizing DEA-based salesperson efficiency benchmarking stems from the use of input and output measures in DEA that do not yield sufficiently discriminating results nor meaningful and actionable directions for improving individual reps' performances based on aspects of their work over which they have direct control. Therefore, in this paper, we propose and demonstrate for the first time how objective metrics for *working hard* and *working smart* – two well-understood concepts in sales management (e.g., Sujan 1986; Sujan, Weitz, and Kumar 1994) – derived from the same archival sales data typically used for sales response analysis or allocation purposes can be fruitfully used as input metrics in DEA-based salesperson efficiency benchmarking. In our application using the proposed working hard and working smart metrics, we show that we obtain better discrimination between efficient and inefficient salespeople than from a conventional DEA that uses simply the raw numbers of sales calls as inputs. Further, we find that *working smarter*, i.e., increasing call effectiveness and/or allocation quality, rather than harder is the more critical direction for improving productivity of the majority of salespeople who were identified as inefficient. The precise extent to which each salesperson should improve in terms of the two aspects of *working smarter* as well as *working harder* was determined by their position and orientation relative to their 100% efficient benchmarks (or reference best-practice set of salespeople). Following these directions to achieve 100% efficiency by all salespeople would yield a gain of more than €13 million in our empirical application.

Of course, we recognize that salespeople should not be evaluated only on the basis of archival sales response data and data-driven benchmarking techniques. A reasonable ‘overall’ benchmarking approach would combine traditional evaluation techniques, like interviews, with data driven evaluation, like the one proposed in this paper.

A limitation of this research is that the proposed model’s fitted benchmarking approach was applied to one medium sized pharmaceutical sales force in our specific application. Other sales force sizes and/or business environments may reveal varying emphases on *working hard* or *working smart*. It would be valuable for sales management to better understand the relative importance of working smarter versus harder in raising productivity in other contexts. Thus, a fruitful direction for further research would be to apply the proposed approach to sales forces from a variety of companies, industries, or countries. We hope our paper stimulates more such research – as sales managers struggle to get more out of their existing sales forces.

Appendix

A. Individualized sales response estimation

We specify a semi-logarithmic sales response function (e.g., Doyle and Saunders 1990) with the observed sales of each salesperson, g per SCU r , and per month t as the dependent variable. The model specification is as follows:

$$Sales_{g,r,t} = \alpha_g + \sum_s \beta_{s,g} \cdot \left[(Pot_s \cdot Docs_{s,r}) \cdot \ln \left(\frac{Calls_{s,r,t}}{Docs_{s,r}} + 1 \right) \right] + \delta_g \cdot PhaCov_r + \gamma_g \cdot t + \varepsilon_g \quad (A1)$$

In Equation (A1) $Sales_{g,r,t}$ are the sales of salesperson g , from SCU r in period t . The independent variable on the right-hand side of (A1) controllable by the salesperson is her/his stock of details $Calls_{s,r,t}$, in month t in the observed SCU r and physician segment s . We use stock variables of the sales calls. For the stock variables, we use a monthly carry-over of 0.9, which has proven to be the best fitting value in a pre-estimation stage. In (A1), the detailing stock is divided by the number of physicians, $Docs_{s,r}$, to obtain the average number of calls per doctor in each segment-SCU combination per month, and segment. Next, to account for uncontrollable potential factors,

we multiply the call variable with SCU-specific prescription potential $Pot_s \cdot Docs_{s,r}$, containing the number of doctors $Docs_{s,r}$ in SCU r and segment s and the segment-specific potential factor Pot_s , which is based on firm data from previous periods.¹³ Another independent variable is an additional potential factor and includes the SCU-specific pharmacy over- or under-coverage $PhaCov_r$, relative to the number of doctors in that SCU. A linear time trend in the (dependent) sales variable is captured by a linear monthly trend t , which has proven to be the best fitting specification. α_g is the salesperson (territory) -specific intercept, and ε_g is the salesperson (territory) -specific error term. We estimate salesperson (territory) g - and segment s -specific detailing coefficients $\beta_{s,g}$, salesperson (territory) specific pharmacy coverage coefficients δ_g , coefficients for the time trend γ_g , intercepts α_g , and error terms ε_g . This specification helps us to account for two sources of heterogeneity. First, we account for observed heterogeneity within individual sales territories and across SCUs by having a SCU-specific potential factor that is multiplied by the individual detailing frequency variables and reflects SCU heterogeneity within the respective territory. Second, we account for unobserved heterogeneity across salespersons (territories) by estimating salesperson-specific individual sales response coefficients. The heterogeneity across SCUs (within territories) is mostly due to the number of doctors and their different distribution across segments, while individual salesperson coefficients reflect the salesperson-specific call effectiveness (working smart). Technically, we use the bayesm package in R (Rossi 2011), a normal prior and a Gibbs Sampler. For estimating the rep level parameters, the first 20,000 draws of a total of 100,000 draws are discarded.

B. Holdout validation of estimated sales response weights using MAPE

Our data represent a panel of 43 monthly periods (2001 to 2004). The first 36 months (2001 to 2003) are used for the individualized response estimation. The remaining seven months in 2004 are used for a holdout validation of the estimated weights. Salesperson-specific response coefficients are used to predict each salesperson's sales volume in each of the remaining seven months of 2004. We evaluate the estimated coefficients by comparing the realized and predicted sales volumes in each of the seven holdout month. Because sales volume data are at the SCU level, and the estimated

¹³ The segment potential multiplier, Pot_s , is the same multiplier which has been used to calculate the segment potential information for the complete territory ($SegPot_{g,s}$) in Equation (1).

sales response coefficients are at the salesperson level (territory level), we aggregate SCU level sales data to the territory level by summing up the SCU level sales volumes for each territory. We calculate the Mean Absolute Percentage Error (MAPE) for each salesperson as the average of seven holdout months. Thus, monthly salesperson level data may be used for the evaluation. Equation (B1) describes the MAPE in formal notation.

$$MAPE = \frac{1}{T} \sum_t \left| \frac{\hat{y}_t - y_t}{y_t} \right| \cdot 100\% \quad (B1)$$

The MAPE calculates the average absolute deviation of the predicted and realized value as a percent of the realized value. t is the currently observed validation period (we use seven monthly periods), y_t is the realized sales volume in t , and \hat{y}_t is the predicted sales volume in t . The average MAPE over all 55 salespersons is 18.30%. The smallest prediction error occurs for salesperson 12 (5.88%), the largest error for salesperson 27 (51.91%), and the standard deviation is 11.75%. Thus, the scale of the MAPE measures implies reasonable estimation quality.

C. The allocation of sales calls on SCU and physician segments

In sales management, allocation periods usually last twelve months. Because we observe data on 36 months, we use three subsequent twelve-month allocation periods, covering the years 2001 to 2003. For calculating the Effort allocation quality according to Equation (2), we use the difference between the optimized and the realized sales volume, averaged across the three allocation periods.

Usually customers are located across several sales coverage units (SCU) inside each territory and several customer potential segments (e.g., VIP, A, B, and C customers). A smart salesperson allocates its calls on those customers, at which her/his individual call has the largest effect on sales. The effect on sales is determined by the specific sales response.

The optimized allocation can be derived from Equation (C1) and (C2). The observed salesperson g has a total budget of sales calls T_g in her/his own territory, which is the sum of a positive number of calls $x_{s,r}$ in physician segment s and SCU r .

$$T_g = \sum_r^R \sum_s^S x_{s,r} \quad (C1)$$

$$Sales_g = \sum_r^R \sum_s^S S_{s,r}(x_{s,r}) \Rightarrow \max! \quad (C2)$$

The individual salesperson's total number of calls T_g is used for generating sales in different physician segments s , and SCUs r , which is expressed by the response function $S_{s,r}$ in Equation (C2). $S_{s,r}$ describes the sales volume in segment s and SCU r generated by $x_{s,r}$ which is the number of calls in physician segment s and SCU r in the observed territory g . The optimized allocation of calls for the observed territory g is found, if the observed allocation of calls $x_{s,r}$ on segments s and SCUs r maximizes the aggregated sales in territory g , $Sales_g$.

Derived from Equations (C2) and (A1), Equation (C3) is a territory sales response function, which is comparable to Equation (A1) but explicitly accounts for the long term effect (marketing multiplier) instead of using stock-variables, and which summarizes sales across SCUs r and segments s on territory level g .

$$Sales_g = \alpha_g + \sum_s^S \sum_r^R \beta_{s,g} \cdot (Pot_s \cdot Docs_{s,r}) \cdot \ln \left(\frac{Calls_{s,r}}{Docs_{s,r}} + 1 \right) + \delta_g \cdot PhaCov_g + \gamma_g \cdot t \quad (C3)$$

$Calls_{s,r}$ is no longer the stock of calls but contains plain calls. The monthly carry-over c (0.9) describes the marketing multiplier. Maximizing Equation (C3) by changing the allocation of calls on SCUs r and segments s (by changing $Calls_{s,r}$) derives the optimized allocation of calls and the resulting optimized sales volume in territory g . The maximization is computed using *Excel Solver*.

D. Bootstrapping for DEA efficiency scores

The DEA models are supplemented by bootstrapped biases for the efficiency scores as described by Simar (1996) and Simar and Wilson (1998, 2000). The bootstrap gives the linear programming results a statistical underpinning and supports the stability of the benchmarking results. In brief the applied method works as follows: We have in

hand for each salesperson g an efficiency score Θ_g estimated using the DEA linear programming algorithm (Charnes, Cooper, and Rhodes 1978; Banker, Charnes, and Cooper 1984). To carry out the bootstrap, we use the following experiment. The data on input factor x_m ($m=1, \dots, M$) for all salespersons, including the observed one, are proportionally scaled using a randomly generated scale factor t (see the papers of Simar and Wilson (1998, 2000), and Simar (1996) for the exact modeling), so that the randomly scaled value RS is calculated as $RS = \Theta_g / t_{m,b}$ for bootstrap replication b ($b=1, \dots, B$) and input factor m . Then, the replicated efficiency score for salesperson g and replication b , $\Theta_{g,b}$, is recomputed using the revised data with the same method. The experiment is repeated B times. The efficiency bias for salesperson g , EB_g is calculated by the difference between the original and the bootstrapped efficiency scores for salesperson g , $EB_g = \Theta_g - \Theta_{g,B}$. In the empirical applications of the recommended and the conventional model we use 300 bootstrap replications ($B=300$).

E. Specification of the conventional DEA input factors

In the conventional benchmarking specification salesperson g 's input factors are the sales calls $Calls_{g,s}$ according to four customer segments s (VIP, A, B, and C). The output is the salespersons' sales volume, $Sales_g$. The *territory potential* again is included as an uncontrollable input factor and calculated according to Equation 4. That is the raw sales response data as used in the proposed benchmarking, however, the latter employs the derived measures *call effectiveness*, the *effort allocation quality*, and the *effort* as proxies for salespersons' *hard* and *smart work*, while the conventional specification uses as input factors the raw factors $Calls_{g,s}$ given by:

$$Calls_{g,s} = \sum_r \frac{1}{T} \sum_t Calls_{r,s,t} \quad (E1)$$

Similar to the sales response estimation described in Equation (A1) $Calls_{r,s,t}$, are modeled as stock of calls using a monthly carryover of 0.9, which has proven to be the best fitting value in a pre-estimation stage. In $Calls_{g,s}$, the calls for SCU r , month t and customer segment s are averaged across month t and summed up across SCU r on territory (salesperson) level g . This aggregation is required by DEA since the method can only make use of cross-sectional data on salesperson (territory) level.

References

- Albers, S. and M. K. Mantrala (2008). Models for Sales Management Decisions, in: Wierenga, B. (Ed.): *Handbook of Marketing Decision Models*, Springer, Berlin et al., 163-210.
- Albers, S., M. K. Mantrala, and S. Sridhar (2010). Personal Selling Elasticities: A Meta-Analysis, *Journal of Marketing Research*, 47 (5), 840–853.
- Avkiran, N. K. (1999). An application reference for data envelopment analysis in branch banking: helping the novice researcher, *International Journal of Bank Marketing*, 17 (5), 206-220.
- Baldwin, H. (2011). Big Change at Big Pharma, *Selling Power*, Cover Story May/June, <http://www.sellingpower.com/content/article.php?a=9459#topArticlePages>.
- Banker, R. D., A. Charnes, and W. W. Cooper (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Management Science*, 30 (9), 1078-1092.
- Boles, J. S., N. Donthu, and R. Lohtia (1995). Salesperson Evaluation Using Relative Performance Efficiency: The Application of Data Envelopment Analysis, *Journal of Personal Selling and Sales Management*, 15 (3), 31-49.
- Brown, S. P. and R. A. Peterson (1994). The effect of effort in sales performance and job satisfaction, *Journal of Marketing*, 58 (2), 70-80.
- Camp, R. C. (1995). *Business Process Benchmarking: Finding and Implementing Best Practices*, ASQC Quality Press, Milwaukee, WI.
- Cegedim-SK&A (2011). *2010 U.S. Pharma Company promotion Spending*, January, http://www.forums.pharma-mkting.com/attachment.php?attachmentid=154&d/2010_pharma_promo_spend_cegedim.pdf.
- Charnes, A., W. W. Cooper, and E. Rhodes (1978). Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2 (6), 429-444.
- Doyle, P. and J. Saunders (1990). Multiproduct Advertising Budgeting, *Marketing Science*, 9 (2), 97-113.

- Dutta, S., W. A. Kamakura, and B. T. Ratchford (2004). Deterministic and Stochastic Approaches for Assessing Technical Efficiency, in: Moorman, C. and D. Lehman (Eds.): *Cool Tools for Strategy: Data Analysis Methods*, Marketing Science Institute.
- Hauser, J. and G. Katz (1998). Metrics: You are what you measure!, *European Management Journal*, 16 (5), 517-528.
- Hershberger, E. K., T. Osmonbekov, and N. Donthu (2001). Benchmarking Marketing Performance, *Working Paper*, 2001, Georgia State University.
- Johnston, M. W. and G. W. Marshall (2010). Salesperson Performance: Behavior, Role, Perceptions, and Satisfaction, in: Churchill, G. A., N. M. Ford, and O. C. Walker (Eds.): *Sales Force Management*, 10th Ed., McGraw-Hill/Irwin, Chicago, 184-213.
- Kamakura, W. A. and T. Lenartowicz, B. T. Ratchford (1996). Productivity Assessment of Multiple Retail Outlets, *Journal of Retailing*, 72 (4), 333-356.
- Ledingham, D., M. Kovac, and H. L. Simon (2006): The New Science of Sales Force Productivity, *Harvard Business Review*, 9, 124-133.
- Lodish, L. M., E. Curtis, M. Ness, and M. K. Simpson (1988): Sales Force Sizing and Deployment Using a Decision Calculus Model at Syntex Laboratories, *Interfaces*, 18 (1), 5-20.
- Mahajan, J. (1991): A data envelopment analytic model for assessing the relative efficiency of the selling function, *European Journal of Operational Research*, 53, 189-205.
- Mantrala, M. K., S. Albers, S. Gopalakrishna, and K. Joseph (2008). Introduction: Special Issue on Enhancing Sales Force Productivity, *Journal of Personal Selling and Sales Management*, 28 (2), 109-113.
- Parsons, L. J. (2004). Measuring performance using stochastic frontier analysis: An industrial salesforce illustration, *ISBM Report*, 6-2004, Institute for the Study of Business Markets, Pennsylvania State University.
- Pettijohn, C., L. S. Pettijohn, A. J. Taylor, and B. D. Keillor. (2001). Are Performance Appraisals a Bureaucratic Exercise of Can They Be Used To Enhance Sales-Force Satisfaction and Commitment?, *Psychology and Marketing*, 18 (4), 337-364.

- Pilling, B., N. Donthu, and S. Hensen (1999): Accounting for the Impact of Territory Characteristics on Sales Performance: Relative Efficiency as a Measure of Salesperson Performance, *Journal of Personal Selling and Sales Management*, 19, 35-45.
- Rackham, N. and J. DeVincentis (1999). *Rethinking the Sales Force: Redefining Selling to Create and Capture Customer*, McGraw-Hill, New York.
- Rapp A., M. Ahearne, J. Mathieu, and N. Schillewaert (2006). The impact of knowledge and empowerment on working smart and working hard: The moderating role of experience, *International Journal of Research in Marketing*, 23 (3), 279-293.
- Rigby, D. and B. Bilodeau (2007). Selecting Management Tools Wisely, *Harvard Business Review*, 85 (12), 20-22.
- Rossi, P. E. (2011). *Bayesm: Bayesian Inference for Marketing/Micro-econometrics*, CRAN, <http://cran.r-project.org/web/packages/bayesm/index.html>.
- Rossi, P. E., G. Allenby, and R. McCulloch (2005). *Bayesian statistics and marketing*. Wiley, Chichester.
- Simar, L. (1996). Aspects of Statistical Analysis in DEA-Type Frontier Models, *Journal of Productivity Analysis*, 7, 177-185.
- Simar, L. and P. W. Wilson (1998). Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models, *Management Science*, 44 (1), 49-61.
- Simar, L. and P. W. Wilson (2000). A general methodology for bootstrapping in non-parametric frontier models, *Journal of Applied Statistics*, 27 (6), 779-802.
- Skiera, B. and S. Albers (1998). COSTA: Contribution Optimizing Sales Territory Alignment. *Marketing Science*, 17 (3), 196-213.
- Spiro, R. L. and B. A. Weitz (1990). Adaptive Selling: Conceptualization, Measurement, and Nomological Validity, *Journal of Marketing Research*, 27 (1), 69-69.
- Sujan, H. (1986). An Exploratory Attributional Analysis of Salespeople's Motivation, *Journal of Marketing Research*, 23 (1), 41-49.
- Sujan, H., B. A. Weitz, and N. Kumar (1994). Learning Orientation, Working Smart, and Effective Selling, *Journal of Marketing*, 58 (3), 39-52.

Sujan H., B. A. Weitz, and M. Sujan (1988). Increasing Sales Productivity by Getting Salespeople to Work Smarter, *Journal of Personal Selling and Sales Management*, 8 (2), 9-19.

Vorhies, D. W. and N. A. Morgan (2005). Benchmarking Marketing Capabilities for Sustainable Competitive Advantage, *Journal of Marketing*, 69 (1), 80-94.

Weitz, B. A., H. Sujan, and M. Sujan (1986). Knowledge, Motivation, and Adaptive Behavior: A Framework for Improving Selling Effectiveness, *Journal of Marketing*, 50 (4), 174-191.

Zairi, M. (1998). *Benchmarking for Best Practice*, Butterworth-Heinemann, Oxford.

Zoltners, A. A., P. Sinha, and S. E. Lorimer (2004). *Sales Force Design for Strategic Advantage*. Palgrave, New York.