Competitive Profiling Displays: Multivariate Graphs for Customer Satisfaction Survey Data

Linda A. Clark, William S. Cleveland, Lorraine Denby and Chuanhai Liu

Linda A. Clark, William S. Cleveland, Lorraine Denby and Chuanhai Liu are Members of the Technical Staff at Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974. The authors would like to thank Judy Gordon, of Lucent's Microelectronics group, for her help in understanding the meaning of the variables and the intricacies of the data and her expertise in competitive surveys. Her knowledge in this area was invaluable.

Competitive Profiling Displays: Multivariate Graphs for Customer Satisfaction Survey Data

Abstract

Competitive customer satisfaction surveys contain a wealth of information for managerial decision making. It is important to convey the analysis of such data in a concise, understandable way. But, typically the results are displayed in tabular form or univariate displays making it difficult to see important relationships in this complex data. This paper describes new multivariate displays that provide a high impact presentation of the survey results, along with their variability, not only for customer satisfaction analysts, but also for upper management.

1 Introduction

In today's competitive market it is critical to understand one's customers. Driving business decisions to meet customer needs and take account of customer opinions is paramount for success. Many instruments are used to obtain these opinions, from report cards, personal interviews and focus groups to competitive studies. Competitive customer satisfaction surveys can provide feedback, not only on one's own customers, but also competitors' customers. These surveys poll customers on satisfaction with important business processes, such as ease of doing business, sales, product quality, ordering, delivery, etc. The choice of these processes depends on the particular industry being studied. Along with these main process areas the respondents may be asked to rate certain subattributes of these processes. For example, subattributes of delivery might be satisfaction with on-time delivery, correctness of the order, etc. In addition, consistent with the conceptual system introduced by Brad Gale (Gale 1994), they typically rate the company on some global variables, such as Overall Satisfaction, Overall Price Satisfaction and Worth What Paid for. Other questions, such as Repurchase or Recommend Intentions may be asked as well.

The quantities of interest are the relative performance of each company to its competitors' performance on these many attributes and subattributes. We define an attribute as a main process area, Overall Satisfaction, Overall Price Satisfaction or Worth What Paid for. These ratios are basic components of customer value analysis (CVA) which claims that the market share of a company is directly linked to the company's performance on the polled attributes (Gale 1994; Naumann and Kordupleski 1995). These ratios, based on both the attribute and subattribute questions, can yield a profile of each competitor across the process areas polled. They are a powerful tool for understanding a company's strengths and weaknesses as compared to its competitors.

Good estimates of the relative performance ratios are a necessary ingredient of understanding customer opinions. But, without careful attention to how these values are displayed, their usefulness and meaning can be diminished. According to Vavra (1997), "One of the biggest downfalls of customer satisfaction measurement is its failure to get properly communicated to operational management. Often the failure is a result of the way in which the information has been presented–it is not user friendly. ... It requires–more than any other type of information–a compelling yet simple recording mechanism." He goes on to say that graphical methods are invaluable for conveying the results. The methods that he describes include unidimensional charts (bar, radar, thermometer, bar and whisker charts) and bivariate charts (competitive leverage and quadrant charts). These are not designed for multivariable data.

Customer satisfaction survey results are often communicated as point estimates without an indication of the associated variability. W. Edwards Deming stated that much of upper management's time is spent chasing spurious variation (Deming 1982). Without the variability and confidence levels of the associated satisfaction scores obtained via a realistic model that takes into account important uncertainties, management and the customer satisfaction analyst are vulnerable to decisions based on random variation.

This paper introduces competitive profiling displays which address both of these issues. They provide a powerful view of each competitor in the industry across all the attributes. Showing the associated variability is an important component of the displays.

These displays will be motivated through the real, though masked, competitive customer survey data described in the next section. The estimates used in the industry and company profiles will be described; but the displays and their interpretation are the focal point of the paper. Several types of competitive profiling displays will be shown.

2 The Data

The data is quarterly survey data, where customers were polled on 9 main attributes and 36 related subattributes. Their responses range from 1 (poor) to 10 (excellent).

The nine main attributes cover quality, technology, price, and interactions between the customer and the company. The attributes and their abbreviated names are:

prod-qual product quality

over-qual overall quality of company processes

delivery delivery of the product

cost cost of the product to the customer, including cost of ownership

features technological excellence of product

pre-sup providing product information before delivery to support customer processes

response responsiveness of the company to customer needs

value value of the product relative to the cost of the product

service service provided by the sales team

The particular subset of data used here covers individuals rating 19 companies over 9 quarters. The total of 3505 pollings reflects 3385 individuals (3271 were polled once, 108 twice, 6 three times). If everyone answered all questions there would be $3505 \times 9 = 31545$ ratings, but skip patterns and missing values bring it down to 25309 observed ratings. We do not analyze the subattribute ratings in this paper.

The survey is designed with a skip pattern such that no respondent rates all nine attributes. Only purchasing agents rate `delivery` and only product and process designers rate `pre-sup`. The remaining seven attributes are intended to be rated by all respondents, but some respondents choose not to rate all attributes about which they are asked. The counts of the number of attributes rated in the 3505 pollings are:

Number of Attributes Rated	1	2	3	4	5	6	7	8
Number of Pollings	5	9	26	64	150	350	1106	1795

The number of individuals surveyed for each of the 19 companies is determined by many factors, including a balance between availability, business profiles and geographic concentration. Since the denominator of the CVA ratio is to reflect the competitive market, sampling proportional to the market share of each company is important. In this data the number of individuals for each company was nearly proportional to its market share.

Table 1 shows the number of ratings for all combinations of attributes and companies. The rows are ordered by their marginal totals increasing from top to bottom. The column marginals increase from left to right.

Due to the sensitive nature of customer satisfaction data, we have masked the names of the companies by renaming them with three letter words such as "Ace," "Bee," "Cab," etc.

3 The Industry and Company Profiles

The industry profile is the relative performance of each company to its competitors with regard to a particular attribute. This set of performance ratings across all competitors enables us to learn how well the industry is performing for that attribute. The company profile is the relative performance of a company across all the attributes measured. Therefore, the quantities of interest are the relative estimates across the nine attribute levels for each of the 19 companies spanning the industry. Estimating the absolute levels for each of these is a necessary component to obtaining the ratios and these performance levels are sometimes of interest themselves.

Obtaining good estimates of these company/attribute levels is a difficult task. Customer survey data is complex due to the skip patterns mentioned in Section 2 and many other sources of missing data: missing at random, sensitive question areas or fatigue factors in answering these questionnaires. There is correlation structure between the attribute areas. Often there are regional differences that need to be taken into account. For example, Europeans may use the scale differently from other regions. Last, but not least, surveys are meant to take opinions in the minds of the customers and turn them into quantitative values. This process is tainted by the characteristics of the raters themselves. An individual may be a hard or an easy grader, may use the entire scale or only part of the scale, or may place different meanings on parts of the scale. The existence of this rater bias is well known (Brady 1985; Brady 1989; Cleveland, Denby, and Liu 1998).

One approach to estimating these levels that takes into account these many varied and complex sources of variability can be found in (Clark et al. 1998). This paper contains a description of an approach which uses Bayesian hierarchical modeling to obtain estimates and their associated variability. This model takes into account (1) the survey respondent effects which are modeled by random location and scale effects, a t-distribution for the location and a Weibull distribution for the scale; (2) company effects for each attribute through time which are modeled by integrated sum-difference processes; (3) error effects which are modeled by a normal distribution, whose variance depends on the attribute.

Clark et al. (1998) found that the data were skewed and suggested the square root transformation, $a\sqrt{(11 - \text{rating})} + b$, scaled to range from 1 to 10, to symmetrize the data. In this paper this transformation was used on the raw survey responses before any analysis; the results are shown on this scale. Any reference to response or rating implies the transformed rating.

Often practitioners calculate performance levels, ignoring these sources of variability, as the average of the untransformed responses for an attribute about a company. These company/attribute estimates are then reported in tables or bar charts, pie charts or line graphs. Typically, no indication of the variability of these values are present (Hayes 1996; Vavra 1997).

Simple estimates for company/attribute values are the $9 \times 19 = 171$ sample means of the transformed data ($\bar{r}_{ac}$ for $c = 1$ to 19, and $a = 1$ to 9). The transformation yields symmetric, but leptokurtic, data. The sample size for each sample mean is large enough that an assumption of normality of sample means is reasonable. Therefore, the expected value of each sample mean is θ_{ac}, the performance level of company c and attribute a. The variance of the sample means is σ^2/n_{ac} where n_{ac} is the number of ratings of company c on attribute a and σ^2 is the variance of a rating about its θ_{ac}.

The sample means are shown in the competitive profiling displays presented in the next section. These estimates are close to the Bayesian estimates obtained in (Clark et al. 1998). As the estimate of σ^2 we use the posterior means of the variances ($\hat{\sigma}_a^2$) obtained in Table 1.3 of (Clark et al. 1998). If one cannot calculate these for the set of data, a simpler estimate is the sample variance s^2 about the company/attribute means. We have fit the model described in (Clark et al. 1998) to two other sets of competitive customer satisfaction survey data and found that these posterior variances described the variability in these other surveys quite well. Therefore, they might be useful in other such data sets.

4 Competitive Profiling Displays

Powerful graphical displays of the industry profiles and company profiles are obtained using the trellis display, which is a framework for the visualization of multivariate data (Becker, Cleveland, and Shyu 1996; Becker and Cleveland 1996). Its most prominent aspect is an overall visual design, reminiscent of a garden trellis, in which panels are laid out in rows, columns, and pages. On each panel of the trellis, a subset of the data is graphed by a display method such as a scatterplot, curve plot, boxplot, 3-D wireframe, normal quantile plot, or dot plot. Each panel shows the relationship of the plotted variables conditional on the values of other variables.

Figure 1 shows the company profiles across each of the attributes surveyed. Each dot is the estimated mean for one attribute and one company. The line segment about the dot shows the 95% confidence interval. Each panel contains the means for each of the nine attributes for a given company. To help in seeing patterns both the companies and attributes are ordered. The attributes are ordered from bottom to top within a panel, according to the means across the companies. The companies are ordered from left to right and then bottom to top through the panels as their means across the nine attributes increase. One enhancement, not shown here, is to plot your company dot (in addition possibly the dot of a competitor of interest) in color or in a different plotting symbol for ease of comparison.

Figure 1 shows that a strong attribute main effect produces similar orderings of the means for the companies. The variation in the values for each company is typically 2 or a bit less. `prod-qual` and `features` are typically largest and second largest, followed by `over-qual`. `cost` is typically smallest and `value` is usually second smallest or close to that position. This rank ordering reflects a natural ordering of the attributes. `features` and `prod-qual` are attributes of physical product. `delivery`, `response`, `service`, and `pre-sup` involve the interaction of the customer with the company personnel. `over-qual` is a mixture of the attributes in these two groups, but falls close to `prod-qual` since it seems to be linked closest to this attribute in the minds of the customers. `cost` and `value` are associated with the pricing of the product and services and the associated value to the customer. The attribute that breaks the ordering most often is `delivery`, but it has the smallest sample size. So, its associated precision is the lowest among the attributes and this larger variation would be expected.

The industry profiles for each attribute in Figure 2 show the same data as Figure 1 but here the estimated levels are graphed against company for a given attribute. Each panel shows the 19 company estimates for one attribute. The companies are ordered bottom to top within a panel so that the company averages across the nine attributes increase. The attributes are ordered so that the one with the smallest average is in the bottom left corner and the largest at the top right. This display shows a notable company effect, though not as strong as the attribute effect; the variation in the values is typically 1 or less. For each attribute, there is a similar ordering of the companies, with the exception of `cost` and possibly `delivery`.

Figures 1 and 2 both show a strong attribute effect. For example, respondents as a group have an industry view that rates `cost` low and `prod-qual` high. Simply looking at an estimate doesn't tell the whole story, because an estimate of 6.5 for `cost` is at the top of the industry, whereas an estimate of 6.5 for `features` is at the bottom. This gives strong evidence for the need for surveying customers of competitors along with one's own. Without the benchmarking provided by the competitor information, it is impossible to evaluate the performance rating of each attribute.

Figures 3 and 4 show the relative company/attribute scores in the same formats as Figures 1 and 2. The plotted value is the company/attribute sample mean shown in Figures 1 and 2 divided by the average of the responses to their competitors for that attribute,

$$\hat{v}_{ac} = \frac{\bar{r}_{ac}}{\bar{r}_{a(-c)}}$$

where

$$\bar{r}_{a(-c)} = \frac{\sum_{j \neq c} n_{aj}\, \bar{r}_{aj}}{\sum_{j \neq c} n_{aj}}.$$

Its variance is approximated by

$$\text{Var}\,(\hat{v}_{ac}) \approx \hat{v}_{ac}^2\,\hat{\sigma}_a^2 \left\{ \frac{1}{n_{ac}\,\bar{r}_{ac}^2} + \frac{1}{\left(\sum_{j\neq c} n_{aj}\right)\bar{r}_{a(-c)}} \right\}.$$

A company whose confidence interval for the ratio does not cross the vertical line at 100 (par) can be considered to be significantly above or below their competitors for that attribute. Figure 3 shows that `Log` is rated significantly above its competition on all attributes except `delivery` where it is at par and `cost` where it is below par. In fact, `Log` has among the highest ratios seen across the industry with a value of 110. On the other hand, `Cab` which sits at the bottom of the industry is significantly below its competitors for all the attributes except `features`, `delivery` and `pre-sup` where the ratios all fall below par but the confidence intervals span par.

Figure 3 gives the industry view for each attribute. Inspecting the `Log` ratios in each panel shows that `Log` has the largest ratio across all the companies for `prod-qual`, `features` (tied with `Duo` and `Nut`), `over-qual` (tied with `Nut`), `pre-sup` and `value`. It is among the top of the industry in `response` and `service`. It is among the bottom of the industry for `cost`. `Cab` is among the worse rated companies for all attributes except `pre-sup`, `delivery`, and `features`.

Businesses commonly set yearly performance goals based on these ratios, possibly even basing compensation on these goals. It is important to set them realistically. Some guidance has been attained through analysis of data for similar companies, sometimes using the PIMS (Profit Impact of Market Strategy) data base (Buzzell and Gale 1987). This analysis can yield a lower bound for "World Class" performance. Since this analysis is not necessarily on direct competitors, it may not give an accurate picture of your industry. Depending on how close one's industry is to the those in the PIMS data base, the "World Class" value may or may not be achievable. Seeing what ratio each competitor in your own industry is achieving provides a better way of obtaining a more realistic performance banding than one based on companies that are not your direct competitors.

5 Time Displays

The competitive profiling displays described in the previous section show performance with all nine quarters rolled together. The means $\bar{r}_{ac}$ are calculated over all available data and, therefore, do not show any trends that might be occurring over the life of the survey.

The trellis display can also be used to show values across time. Suppose that the quarterly sample relative means are $\hat{v}_{acq}$ of attribute a for company c in quarter q. Figure 5 is the first page of a five page trellis display showing the relative quarterly levels for the companies with the most data. It is clear that the quarter to quarter variation is large and spurious and can mask any trends that might be present. The relative performance probably moves smoothly through time; no large meaningful large quarter-to-quarter jumps would be expected. In this display the spurious variation masked any trends.

To make it easier to identify any smooth trends, the quarterly values are smoothed and displayed. Figure 6 shows the loess smooths of the adjusted ratings against quarter a given company and attribute (Cleveland and Loader 1996). The curve on given panel is a smooth of the absolute quarterly levels of the attribute for the company across time divided by the smooth of the competitive level. The loess smoothing parameter is 1 and local linear fitting is used with weights proportional to $1/Var(\bar{r}_{acq})$ for the numerator and $1/Var(\bar{r}_{a(-c)q})$ for the denominator. The variance of the ratio is computed using the variance of the weighted smoothed values for the numerator and denominator.

Clearly, the quarterly effects are small. These changes are low frequency effects. There are no convincing short-term effects such as jumps in level. Any time changes that might be present have small magnitudes compared with the changes in the overall levels between the companies.

Suppose an attribute for a company whose relative performance seems to be improving or declining (e.g., `delivery` for `Fan`). Then one might want to question whether `Fan`'s relative performance is declining because `Fan` is decaying or its competitors' performance is improving. Figure 7 shows a plot that can answer this question for `Fan` and `Duo` for all nine attributes. This display shows the smooths for both `Fan`'s `delivery` and for its competitors' `delivery` performance. In Figure 6 the relative performance of `Fan`'s `delivery` is declining. From Figure 7 it is clear that this is due to the fact that `Fan`'s `delivery` score (the solid line) is actually declining and the competitive level (the dashed line) is remaining constant.

6 Conclusions

The competitive profile displays and time displays described here are very powerful for conveying complex customer survey data. They not only enable one to understand large sets of data but lead to more insights than the typical tabular summaries that are often used. Further, it is very important to give an assessment of precision when presenting customer satisfaction results to avoid making business decisions based on random variation.

These displays have been presented to customer satisfaction managers and senior executives at Lucent Technologies. Both of these audiences have found them to be far more useful than previous presentation methods. They are currently being adopted for displaying customer satisfaction results in other areas of the company.

	deli	pre-	prod	serv	cost	resp	valu	feat	over
Nut	15	15	25	29	28	32	33	33	32
Gas	17	25	26	34	40	43	42	44	46
Ham	13	35	38	43	42	47	46	50	50
Pub	32	32	56	59	61	67	68	70	71
Key	35	41	49	70	72	76	77	77	76
Cab	25	74	65	85	88	104	107	113	111
Ear	28	74	75	92	94	102	106	106	107
Bee	38	76	87	114	119	121	128	131	132
Mug	58	71	106	116	120	132	136	138	137
Jet	26	128	102	138	147	153	158	162	162
Rug	54	100	124	148	163	169	174	167	174
Log	63	97	121	156	165	167	170	172	171
Toy	47	110	132	145	152	170	177	178	181
Oak	84	110	169	174	197	205	218	219	216
Ace	77	174	186	254	240	270	267	273	275
Inn	83	187	180	255	260	283	287	287	285
Ski	112	189	250	282	300	321	334	328	335
Duo	121	292	309	408	389	420	420	431	430
Fan	132	303	320	400	403	431	455	465	463

Table 1

Counts of Pollings According to Number of Questions Answered

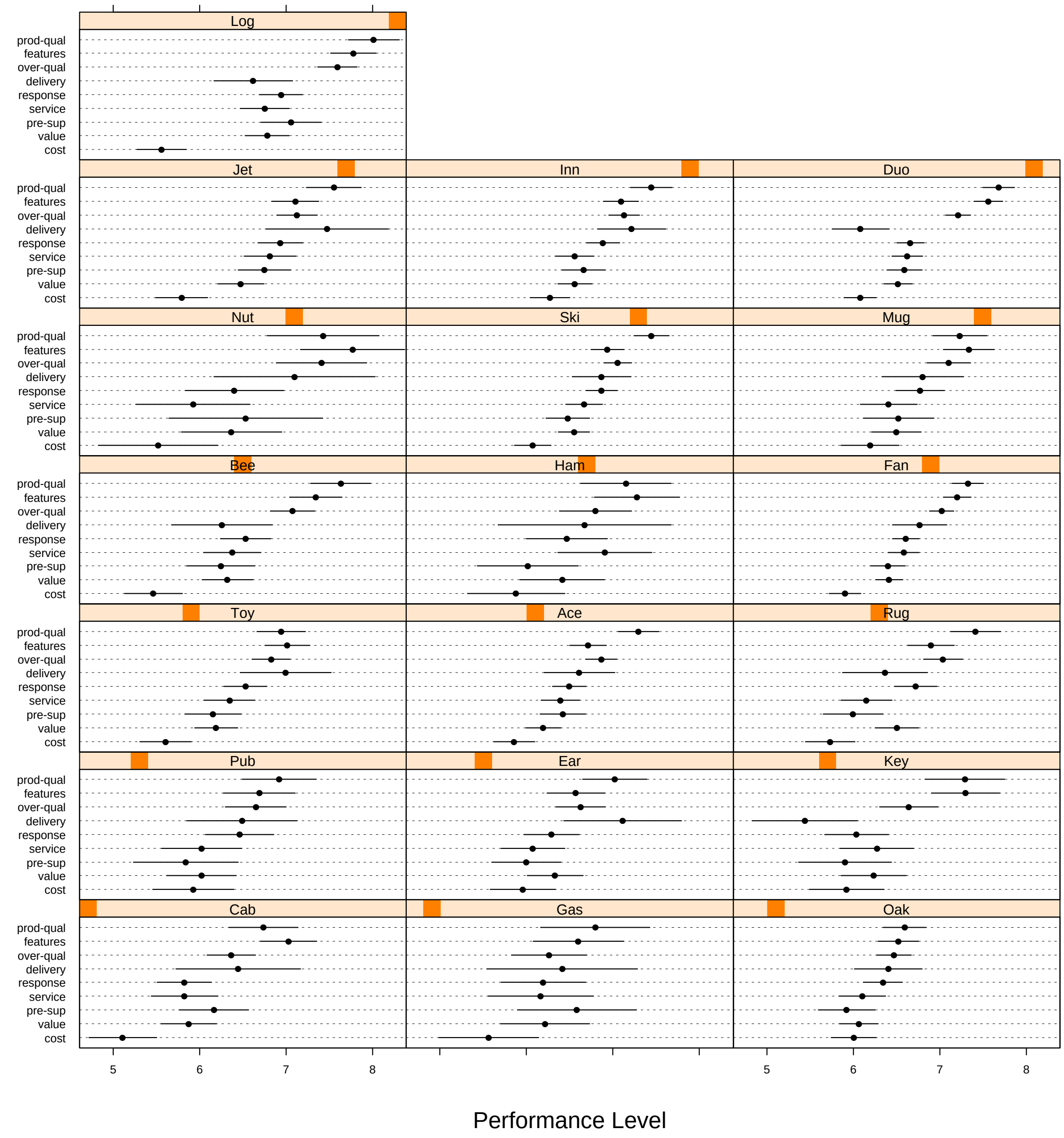

Figure 1
Nine-Quarter Average Rating Given Company

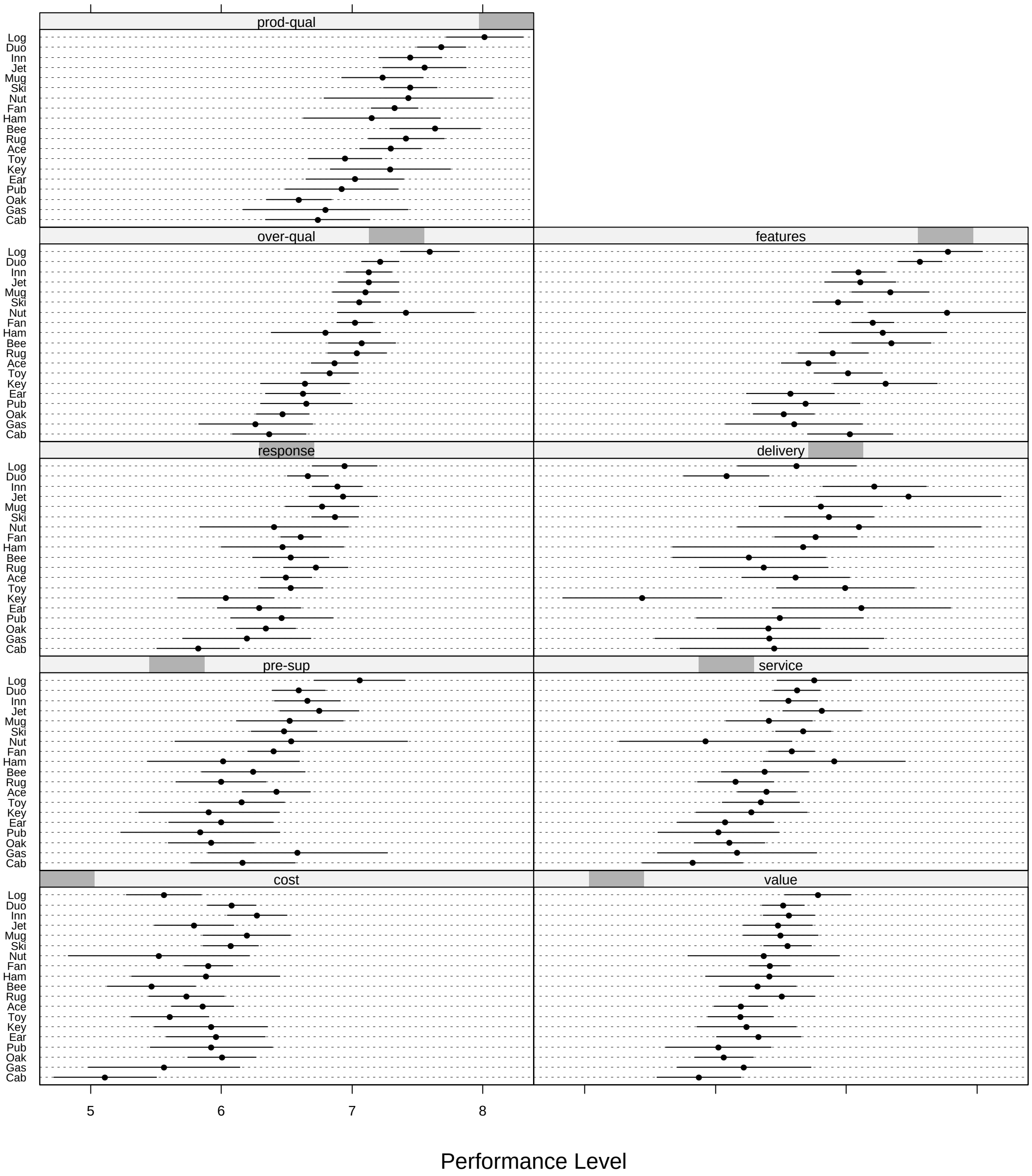

Figure 2
Nine-Quarter Average Rating Given Attribute

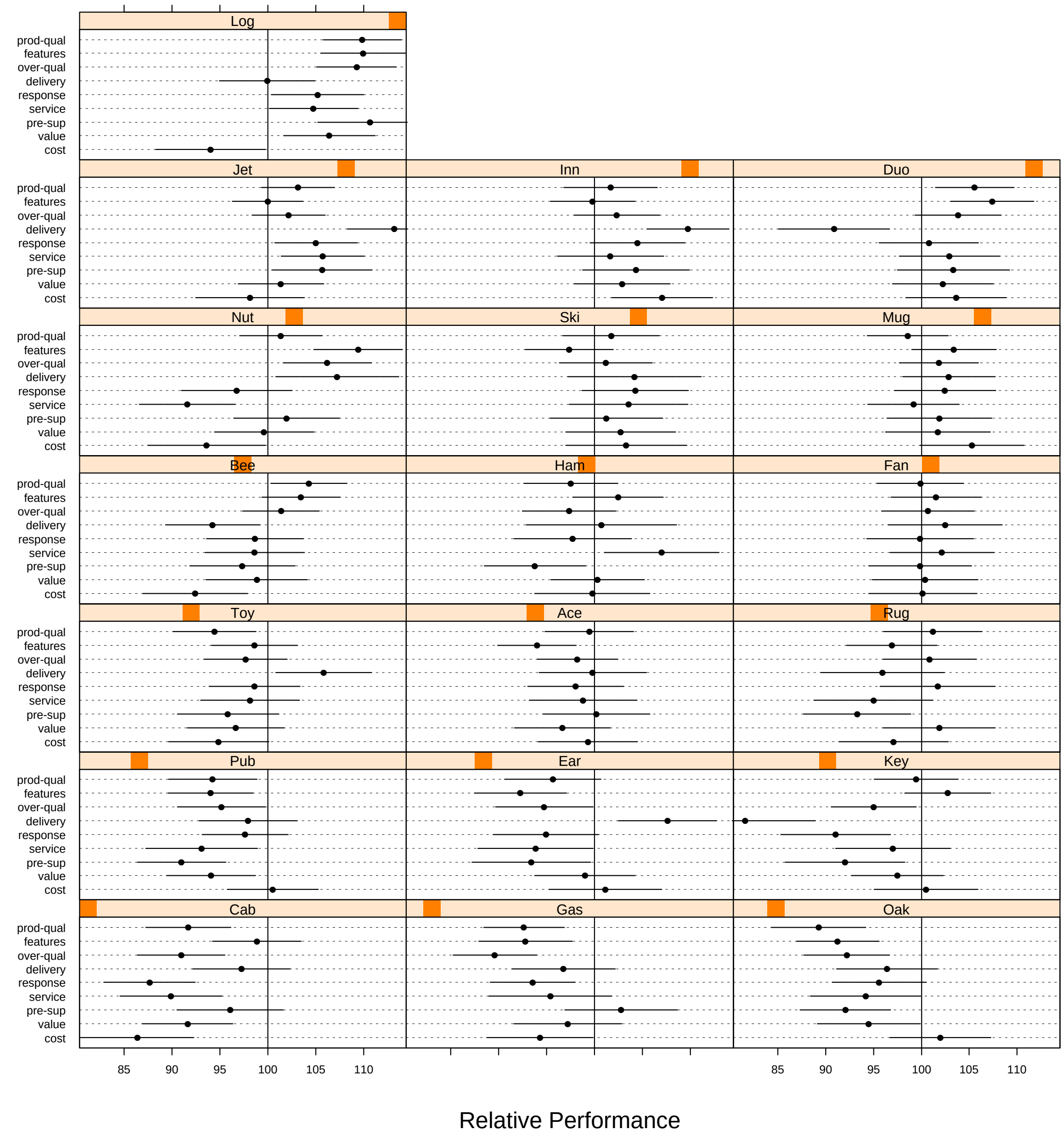

Figure 3
Nine-Quarter Relative Performance Given Company

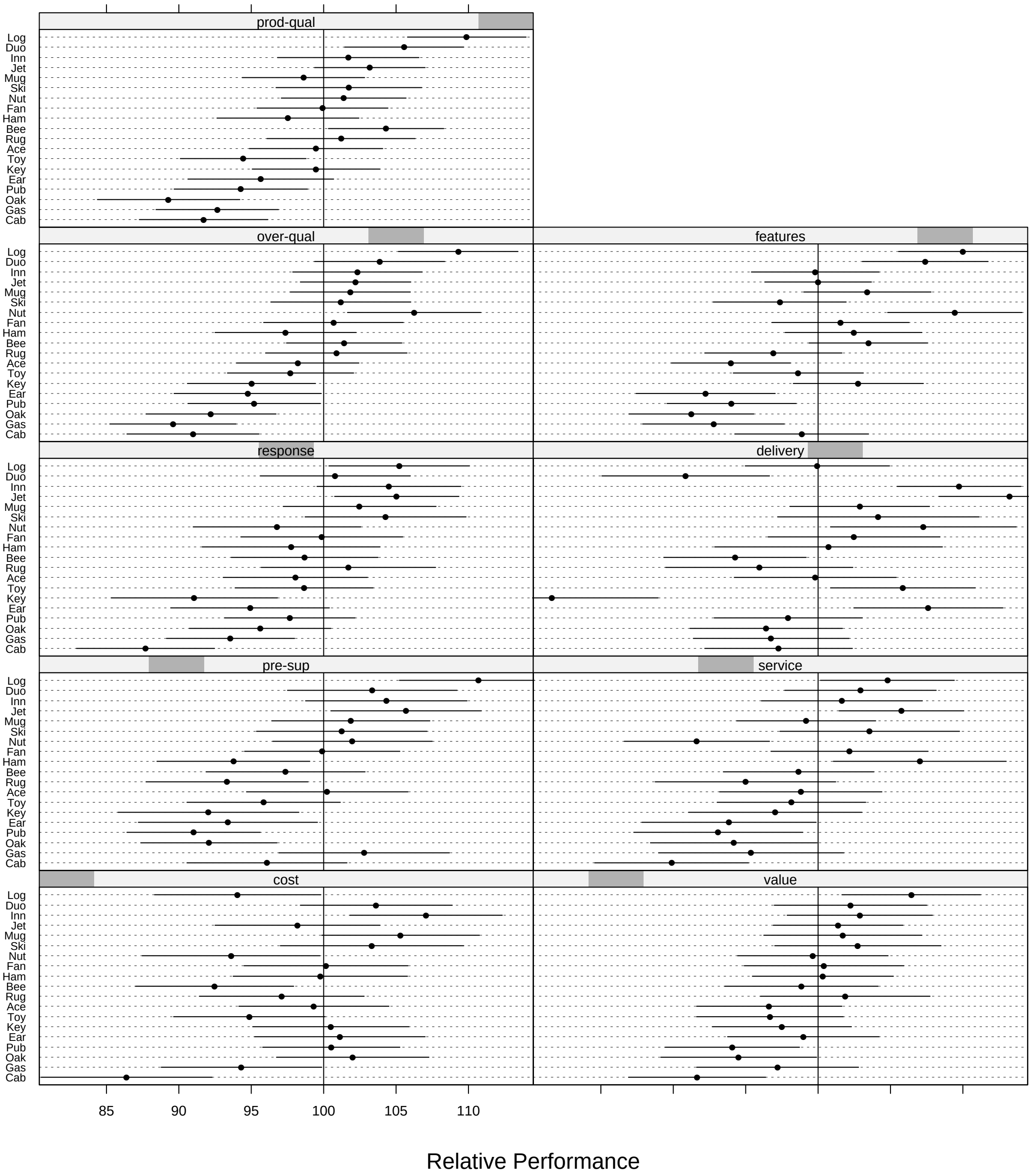

Figure 4
Nine-Quarter Relative Performance Given Attribute

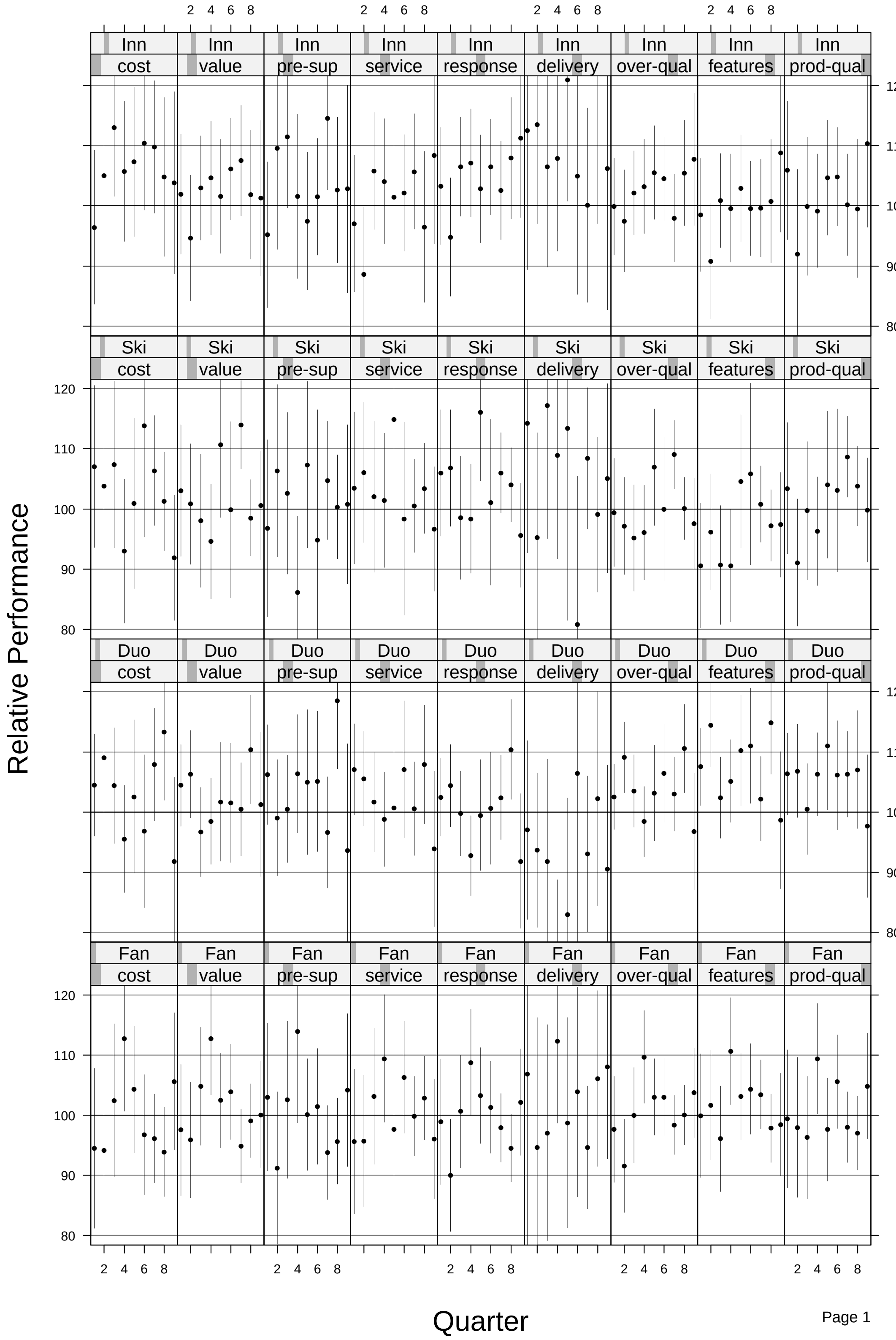

Figure 5
Quarterly Relative Performance Given Attribute and Company

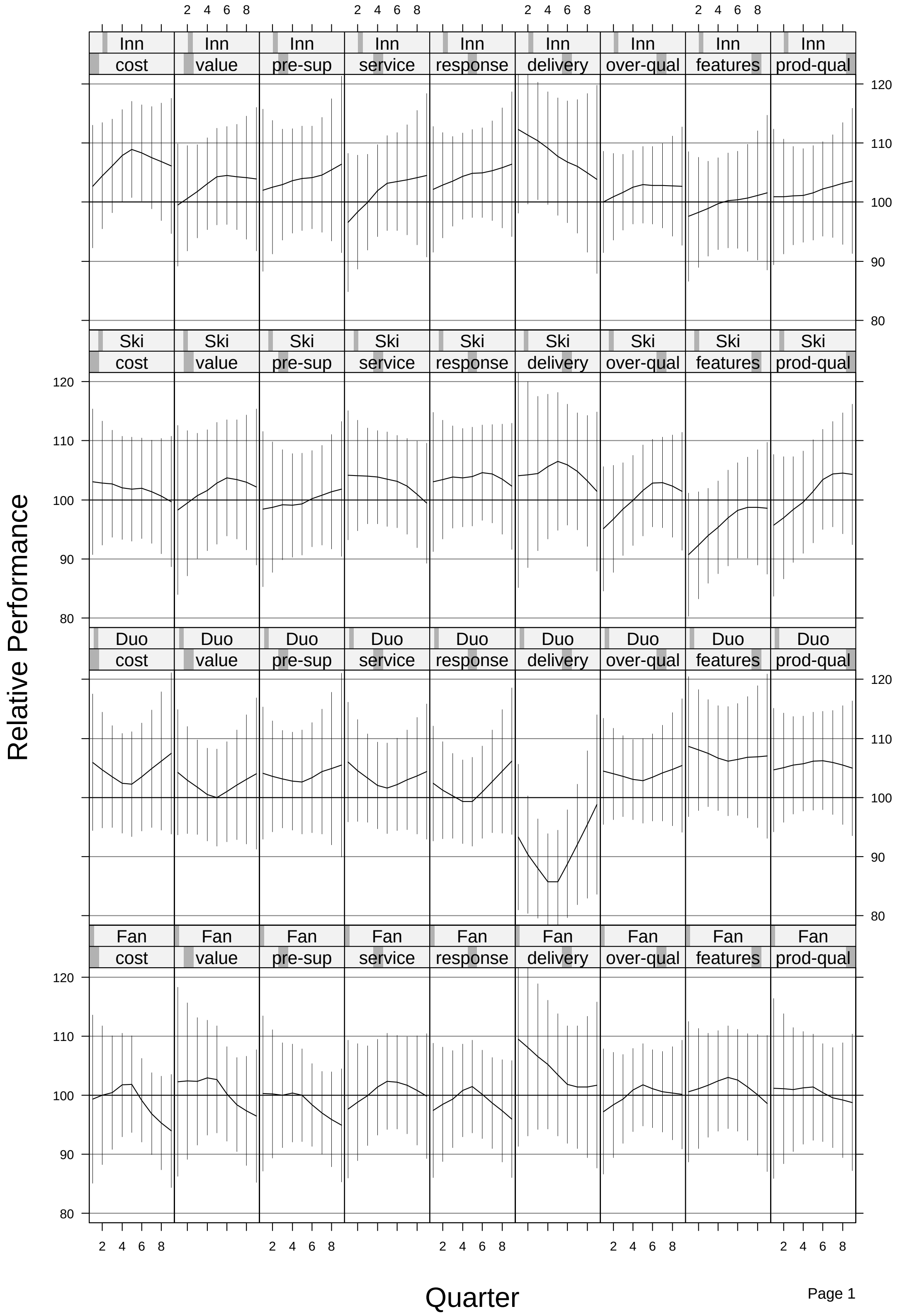

Figure 6
Smoothed Quarterly Relative Performance Given Attribute and Company

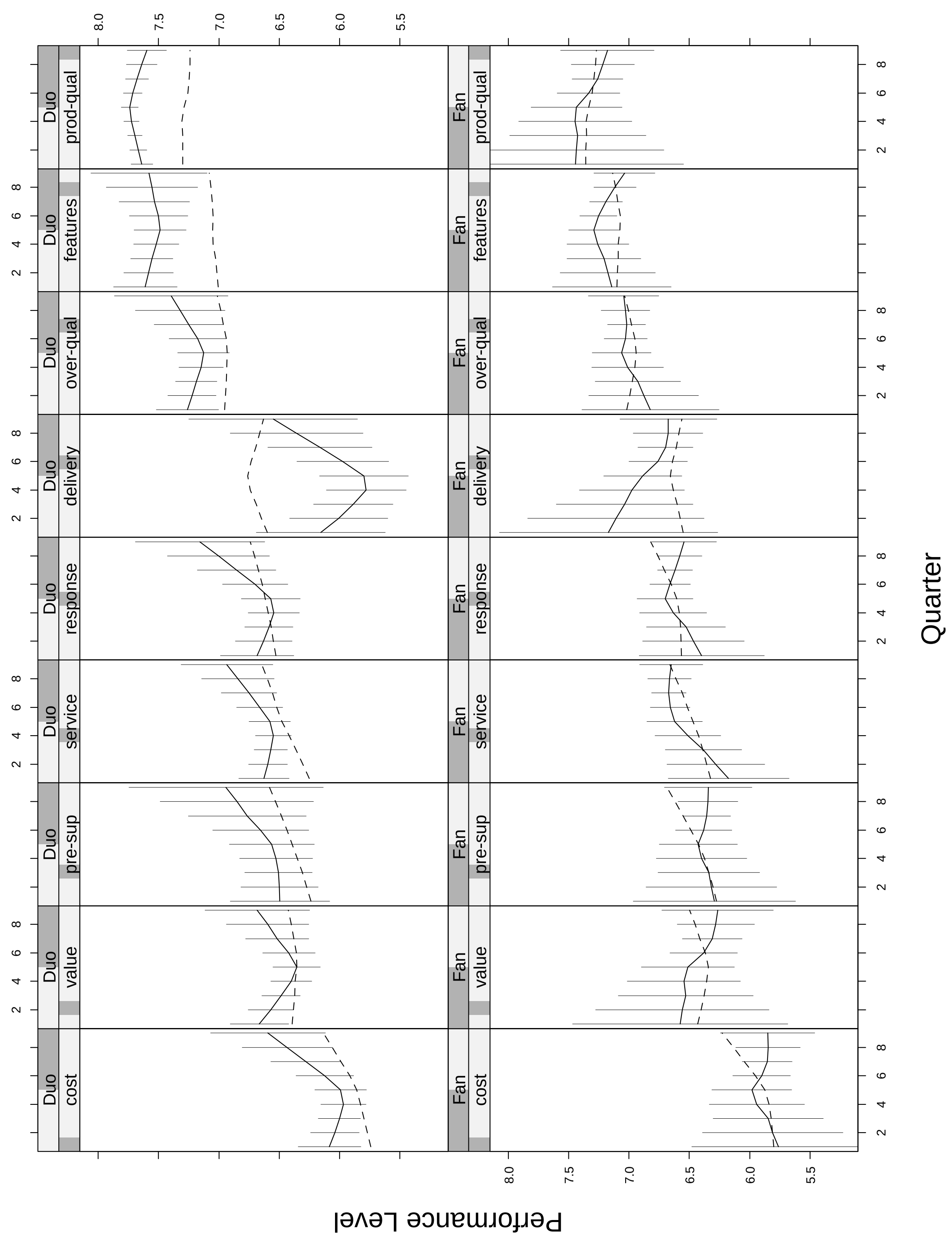

Figure 7

Smoothed Quarterly Performance for Two Companies and Their Competitors Given Attribute

References

Becker, Richard A. and William S. Cleveland (1996), *Trellis Graphics User's Manual*. Seattle, WA: MathSoft.

B&W postscript — http://cm.bell-labs.com/stat/doc/trellis.user.bw.ps.

Color postscript — http://cm.bell-labs.com/stat/doc/trellis.user.color.ps.

Becker, Richard A., William S. Cleveland, and Ming J. Shyu (1996), "The Design and Control of Trellis Display," *Journal of Computational and Statistical Graphics*, 5, 123–155.

B&W postscript — http://cm.bell-labs.com/stat/doc/trellis.jcgs.bw.ps.

Color postscript — http://cm.bell-labs.com/stat/doc/trellis.jcgs.col.ps.

Brady, Henry (1985), "The Perils of Survey Research: Interpersonally Incomparable Data," *Political Methodology*, 11, 269–292.

— (1989), "Factor and Ideal Point Analysis for Interpersonally Incomparable Data," *Psychometrika*, 50, 509–537.

Buzzell, Robert D. and Bradley T. Gale (1987), *The PIMS Principles: Linking Strategy to Performance*. New York: Free Press.

Clark, Linda A., William S. Cleveland, Lorraine Denby, and Chuanhai Liu (1998), "Modeling Customer Survey Data," in *Case Studies in Bayesian Statistics IV*, C. Gatsonis et al., eds. New York: Springer.

Cleveland, William S., Lorraine Denby, and Chuanhai Liu (1998), "Random Location and Scale Models and Model-Building Methods for Case Data," (in preparation).

Cleveland, William S. and Clive Loader (1996), "Smoothing by Local Regression: Principles and Methods," in *Statistical Theory and Computational Aspects of Smoothing*, W. Hardle and M. G. Schimek, eds. Heidelberg: Physica Verlag.

Deming, W. Edwards (1982), *Out of the Crisis.* Massachusetts Institute of Technology Center for Advanced Engineering Study. Cambridge, MA.

Gale, Bradley T. (1994), *Managing Customer Value.* New York: MacMillan.

Hayes, Bob E. (1996), *Measuring Customer Satisfaction: Survey Design, Use & Statistical Analysis Methods*, Second Edition. Milwaukee, Wisconsin: ASQC Quality Press.

Naumann, Earl and Ray Kordupleski (1995), *Customer Value Toolkit.* London: International Thomson Publishing.

Vavra, Terry G. (1997), *Improving Your Measurement of Customer Satisfaction: A Guide to Creating, Conducting, Analyzing, and Reporting Customer Satisfaction Measurement Programs.* Milwaukee, Wisconsin: ASQC Quality Press.