BEST PRACTICE APPLICATION: IDENTIFYING HIGH AND LOW BEHAVIOR AND PERFORMANCE USING QUANTILE REGRESSION AND SWLS MODELING

Jiannan Wu and Stuart Bretschneider

ABSTRACT

How can we identify best-practice providers? Under the combined influence of GPRA[1], the NPR, the state and community benchmarking efforts, and GASB SEA reporting requirements, most federal, state, and local government agencies, private for-profit and nonprofit organizations delivering government programs under grants and contracts, will become involved in performance measurement. Once governments begin routinely collecting and reporting performance measurement data, policymakers and policy evaluators will be faced with the task of identifying best-practice providers. How can governments go about making comparisons among service providers using performance measurement data? Can best-practice providers actually be identified? Based on previous analysis using a Quantile Regression and SWLS model for estimation and inference, this article introduces a new approach to estimating models of extreme behavior. Quantile Regression and SWLS are investigated to lay a foundation for putting forward the new analysis technique: Segmentation Strategy. Then, some preparatory work for Monte Carlo Simulation, including determining the structure of simulated data sets, is described. Thirdly, the computational results are displayed and analyzed. Finally, some conclusions and future research directions are provided.

INTRODUCTION

Most of the statistical tools for theory estimation and validation focus on what may be termed typical behavior in management. However, in some settings we are more interested in the extremes. For example, in the management field, best practice has had a long established and recognized track record. From Frederick Taylor Scientific Management to Elton Mayo's Human Relation Movement, from Peters and Watermans' 'Search of Excellence' to Osborne and Gaebler's 'Reinventing Government', all of those management gurus and their work have used the thoughts and methods of best practice (Katorobo, 1998; Overman and Boyd, 1994). In the practice of public affairs, interests in best practice have been visible at all levels of government. For example, the state of New York has established a standing committee on Best Practices. The National Governor's Association, the General Services Administration, and numerous other organizations have established awards for best practices involving IT systems and projects (Rocheleau, 2000). Actually, the issue of performance measurement, the regular collection and reporting of information about the efficiency, quality, and effectiveness of government programs, is arguably the hottest topic in government since 1990s (Nyhan and Martin, 1999).

However, how can we identify the best-practice providers? The combined influence of GPRA[2], the NPR, the state and community benchmarking efforts, and the GASB's SEA reporting virtually ensures that most federal, state, and local government agencies, as well as those private for-profit and nonprofit organizations delivering government programs under grants and contracts, will become involved in performance measurement before the end of decade (Nyhan, and Martin, 1999). Once governments begin routinely collecting and reporting performance measurement data, policymakers and policy evaluators will be faced with the task of identifying best-practice providers. How can governments go about making comparisons among service providers using performance measurement data? Can best-practice providers actually be identified? Just as Nyhan and Martin (1999) said, “after encouraging governments to routinely collect and report information on the efficiency, quality, and effectiveness of their programs, the performance measurement literature suddenly becomes silent on how the resulting data might be used in making service provider comparisons”. A similar problem also exists in information management of public sector organizations, where qualitative judgment of experts has been the primary source of identifying best practices and (Rocheleau, 2000).

There exist several approaches to identify the extreme performers, including the best or worst performers. Data Envelopment Analysis (DEA) has been used to locate the best provider by estimating a measure of relative technical productive (Charnes Cooper and Rhodes, 1978). Quantile regression is another method which will provide statistical estimates of conditional relationships for extreme cases (Marc-Arulle, D’Amico and Bretschneider, 2000). This method has been used to directly estimate models for upper quantiles of conditional distribution rather than inferring such relationships based on conditional central tendency (Bassett and Koenker, 1978, 1982; Koenker and Basset, 1978; Eide and Showalter, 1999; Scharf, Juanes, and Sutherland, 1998; Koenker, 2000b). Another method called Substantively Weighted Least Squares (SWLS)[3] has also been introduced and applied in the field of public administration and policy to investigate optimal performers (Meier and Keiser, 1996; Meier and Gill, 2000). However, those researches only provide limited understanding of how these techniques will actually work in the real world.

Consequently, recent work in this area has made use of Monte Carlo studies to better understand how well these techniques uncover the true state of nature when extreme organizations differ from typical ones. Wu et. al. and Marc-Arulle et. al. have compared the estimation performance of extreme behavior between SWAT and Quantile Regression using one underlying case and one hundred simulated samples to demonstrate that Quantile Regression had greater efficiency at estimating extreme behavior (Marc-Arulle, D’Amico and Bretschneider, 2000; Wu, Bretschneider, and Marc-Arulle, 2000). This article will introduce a new approach to estimating models of extreme behavior and study how well it compares to Quantile Regression and SWAT in a wide variety of simulated situations. The next section of the paper provides the theoretical foundations for Quantile Regression and SWLS as well as the basis for our new approach based on a Segmentation Strategy. This is followed by a discussion of our design for a Monte Carlo Simulation, including determining the structure of simulated datasets. Next we present the computational results from the Monte Carlo simulations. Finally, some conclusions and next research steps are given.

ANALYTICAL TOOLS FOR EXTREME BEHAVIORS: MODEL, ESTIMATION AND INFERENCE

To understand any analytical tools in statistics, the underlying model, estimation process and inference procedures are fundamental. We next address each of these issues for Quantile Regression, SWAT and our proposed Segmentation Strategy.

Quantile regression

The model

Quantile Regression is a statistical technique intended to estimate, and conduct inference about, conditional quantile functions (Koenker, 2000a). Its model, first introduced by Koener and Bassett (1978b), can be written as

$$y_i = x_i^{'}\beta_\theta + u_{\theta_i}, \qquad Quant_\theta(y_i \mid x_i) = x_i^{'}\beta_\theta \qquad (1)$$

where (y_i, x_i), $i = 1,...,n$, is a sample of n cases from some population, x_i is a $K \times 1$ vector of regressors, and $Quant_\theta(y_i \mid x_i)$ denotes the conditional quantile of y_i, conditional on the regressor vector x_i (Buchinsky, 1998; Eide and Showalter, 1999). Note that as θ increases continuously from 0 to 1, it traces the entire conditional distribution of y conditional on x. Given the practical situations of dataset, there are always a number of quantile estimates which will be numerically distinct though this number is typically unknown. More interestingly, there might be different groups of observation whose quantile estimates might be significant. Based on these considerations, we rewrite the model as

$$Y = \begin{cases} a_1 + b_1 X_1 + \varepsilon_1 \\ a_2 + b_2 X_2 + \varepsilon_2 \\ \dots \\ a_m + b_m X_m + \varepsilon_m \end{cases} \qquad (2)$$

From Equation (2), we can see that quantile regression may generate different models at different location. When m=1, there is only one model hidden in the datasets, the traditional conditional mean or median case. When m=3, there exist three distinctive model structures, which might represent the best, moderate and worst situations. When m>3, we must readjust our thinking to include more subgroups and structures. By isolating the different groups hidden in the dataset, we can estimate how these units performers and how they differ from each other.

Estimation and issues

Koenker and Bassett (1978, 1982) developed the theory for estimation of the conditional quantiles of a variable y_t that is assumed to be a linear function of other variables. Later Bloomfield and Steiger (1980) and Koenker and D'Orey (1987) detail the problem as on of minimizing a sum of absolute deviations as a linear programming problem, which is the basis for estimation of the model parameters.

Specifically, the estimation of Quantile Regression is done by minimizing expression (3):

$$\sum_{t\in\{t:y_t\geq b\}}^{n} \theta \mid y_t - x_t\beta \mid + \sum_{t\in\{t:y_t<b\}} (1-\theta) \mid y_t - x_t\beta \mid \qquad (3)$$

Where y_t is the dependent variable, x_t is the explanatory variable, θ is the quantile to be given, and β is the estimate of b. When $\theta=0.5$, the solution will be identical to the estimates produced by Least Absolute Estimation (LAE) technique. Quantile regression had been used in a broad range of application settings as a comprehensive approach to statistical analysis of linear and nonlinear response models (Buchinsky, 1998; Koenker, 2000a and 2000b).

Inference and issues

One reason for reluctance of researchers to use median regression (LAE), one of special case of quantile regression, had been the absence of any theory to provide for inference about the model parameters. In 1978 Basset and Koenker (1978) developed a formal basis for large sample inference for the special case of the median regression and later extended it to include the full set of quantile regression models (Basset and Koenker, 1982). Most of the early simulation studies of LAV estimators were to compare the small sample efficiency of LAV and least squares estimators for various error distributions. The results from those studies indicated that LAV estimators were more efficient than least squares estimators. The results for large samples implies that for any error distribution for which the median is (asymptotically) superior to the mean as an estimator of location the same holds for LAV over OLS regression. In other words, LAV estimation is preferable to least squares (Dielman and Pfaffenberger, 1982) in many situations.

In sum, Quantile Regression applies a multi-scope model that differentiates subgroups within a conditional distribution of the dependent variable. Its estimation is realized by minimizing the conditional quantile function based on solving a linear programming problem. Statistical inference on parameters for either small sample or large sample tends to be more efficient thaner than mean-based regression in many real situations, especially when data have outliers and asymmetric distributions. Most importantly, by selecting different ‘locations’, either in the low or high ends of the distribution, extreme behaviors, either best or worst, can be located objectively. Computation implementation is relatively easy using a variety of available statistical software.

SWLS

In the analysis of the performance of a state level program for child support across 50 states, Meier and Keiser (1996) tried to find which factors have the greatest impact on successful implementation. In this work they proposed the Substantively Weighted Least Squares (SWLS) method, in which outliers, i. e. extreme observations (Neter, et al, 1996), are regarded as potential prescriptions for improving future performance (Gill, 1997).

The model

Interested in high-performing agencies, Meier and Keiser (1996) overweighed the extreme cases rather than down weighting them as robust regression analysis does. Specifically, they used a studentized residual of 0.7 as a threshold of high performance. All cases from an initial OLS regression with "studentized" residuals of 0.7 or less were down weighted, the average agencies, in a series of iterated re-estimation by increments of 0.1 until the average cases were all equally weighted as only 0.1 to the 1.0 for high-performing cases. By comparing the change of slope coefficient in each of the re-weighted regressions, the analyst could see what the high-performing agencies did that the average ones did not.

SWLS regard observations in the dataset in two ways: those with studentized residuals of higher than 0.7 and those less than 0.7. Each observation in the dataset, either belongs to the better performing group, or belongs to the remaining group. SWLS procedures have been implemented based on a -0.7 studentized residual to successively partition the lower end of a dataset. In this fashion SWLS attempts to identify low performing behaviors relative to the typical weight of 1.0 to all sample cases with initial "studentized" residuals above the -0.7 partitions. Hence, the model of SWLS can be expressed as follows:

$$Y = \begin{cases} a_1 + b_1 X_1 + \varepsilon_1 \\ a_2 + b_2 X_2 + \varepsilon_2 \end{cases} \qquad (4)$$

In subsequent applications of SWLS the authors permit the magnitude of each re-weighted iteration to be adjusted to values other than 0.7 or -0.7 to form what the authors refer to as a generalized version of the process called Substantively Weighted Analytical Technique (SWAT). Of course, with the change of 'substantive weight', the estimates will change. And, with the change of threshold of high performance, the estimates will change, as well. The problem of *apriori* setting the weighting criteria makes the process heuristic and adds a significant judgmental component to the process.

Estimation and issues

Suppose that we fit a local regression to the data, obtaining estimates $\hat{y}_i$, residuals $e_i = y_i - \hat{y}_i$, studentized residuals $t_i = e_i / s_i$ 4, and weights $W_i = W(t_i)$ for each observation. The estimation of SWLS is to be obtained by minimizing the exponential term:

$$Q_W = \sum_{i=1}^{n} W_i (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_{p-1} x_{i,p-11})^2 \qquad (5)$$

From a mathematical perspective, this process is similar to the techniques of Robust Regression (Huber 1973). The distinction is that robust regression weakens the influences of extreme cases by giving them less weight and embodies the situations of average performances while SWLS innovatively place more weight on extreme observations. Moreover, in the execution of SWLS, analysts have to subjectively determine the value of the weight and the number of iterations, i. e. the arbitrariness of the jackkniefed residuals and the number of down-weighting iterations (Gill, 1997).

Inference and issues

As a technique for performance isolation and recommendation, SWAT has been used to investigate optimal performers (Meier and Keiser 1996), multiple goals (Meier, Wringkle, and Polinard 1999a), risk averse and failing organizations (Meier, Gill, and Waller 2000), minority representation (Meier, Wrinkle, and Polinard 1999b), and the differences between good agencies and exceptional ones (Gill and Meier forthcoming). However, we have no sampling theory to base any statistical inference on when using SWLS.

According to the developers of SWLS, it is a new approach that they feel "has the potential to transform the basic quantitative method of public administration (regression) from a tool that explains what is to a tool that can be used for search for what might be". And, as they emphasized in their articles, SWLS or SWAT is applied as a ‘Qualitative’ method and it can tell the analyst ‘what might be’. Recently, in their book titled ‘What works: A new approach to Program and Policy Analysis’, the authors of SWLS/SWAT admitted that SWAT results cannot say anything about whether or not S (Sample) could be drawn from P (Population) since SWAT reweighs outlier cases (Meier and Keiser, 2000).

In our opinions, since SWLS or SWAT borrows heavily for least squares regression methods to arrive at its conclusions, it must rely heavily on the assumptions of least squares regression—thus potentially creating some weaknesses in the analysis of extremes. This clearly suggests that before one can prescribe the use of a technique we must have a greater understanding of how it will operate and perform under know conditions before applying it to situations where the underlying structure is unknown. Axiomatic theory clearly is insufficient to guide us in selecting an empirical approach. This strongly suggests the need for comparative analysis in simulated situation before applying these approaches.

Prior Comparative studies

Two prior studies have attempted to compare SWLS/SWAT with quantile regression using a simulation (Marc-Aurele, D' Amico and Bretschneider, 2000; Wu, Bretschneider, Marc-Aurele, 2000). These studies while useful provided only a limited basis for comparison relying on a simulated case where the numbers of high, low and medium performing units were equal and relatively large in number (e.g. n=200). However, a stricter and more systematic framework of analysis is still needed to further understand how those tools will perform in a number of real world situations.

Segmentation Strategy

The above analysis of Quantile Regression and SWLS, suggests that both approaches rely on an underlying model has similar structures, which suggests a more general approach. Towards that end we developed a third approach based on the idea that regardless of how the estimation is done, the model must be based on some form of *apriori* segmentation of the data into some number of m groups.

The model

The Quantile regression model structure was based on a series of up to 100 separate regression quantiles, though for most practical applications three groups tend to be sufficient; high, middle and low. The SWLS approach divided the data into two groups, the high (or low) and everything else. These two ideas can be generalized into the idea that some unknown number of groups exist in the data. Each of these groups is a segment thus we propose a general model of Segmentation where the underlying structure has the form of equation (6).

$$Y = \begin{cases} \alpha_1 + \beta_1 X_1 + \varepsilon_1 \\ \alpha_2 + \beta_2 X \quad + \varepsilon_2 \\ \ldots \\ \alpha_n + \beta_n X_n + \varepsilon \end{cases} \qquad (6)$$

In practical analysis, it is necessary to adopt some methods for determining n, how many models are hidden within the dataset. Cluster Analysis or Data Mining may be able to solve this problem though, to date there are no satisfactory methods for determining the number of population clusters for any type of cluster analysis (Everitt, 1979, 1980; Hartigan 1985; Bock 1985; SAS/STAT User's Guide, 1990, p97). Another more heuristic approach is to look at the data graphically and apply various forms of exploratory data analysis to induce the number of substructures in the data (SAS/STAT User's Guide, 1990, p97).

Estimation and issues

Once the number of subgroups or segments in the dataset is defined, any standard estimation tool can be use including least squares, robust or least absolute deviation regression can be run to estimate unique models for each segment. One specific approach is a modification of the SWLS approach by forming three groups of cases (the typical situation for high, middle, and low performers) by sorting cases based on the "studentized" residuals from an initial least squares regression. Concretely, if the number of segments is three, we will propose that cases with studentized residual of +.7 or more be designated as the high-performing cases, "studentized" residual of -.7 or less as the low-performing cases and those between 0.7 and -0.7 be included in the middle group. Once identified three separate least squares models can be run on the respective three groups. The results of the three OLS regressions will tell me how the three groups operate and identify the best, middle and worse cases.

Inference and issues

Since the essence of the Segmentation approach is to first define a set of n separate samples within a single dataset, inference is possible for each of the n separate estimated models based on the technique of estimation applied. If one uses least squares estimation on the individual model standard inference applies assuming the sub-sample meet s the typical sampling requirements. Similarly, if the sub-models are estimated using robust or median regression, the theory of inference that applies to those models would be applicable.

The essence of the approach is to *apriori* identify either analytically of qualitatively the number of segments in the dataset and sort the data into groups of cases so that individual and independent sub-models can be estimated. To simply this process we suggest two heuristics, the first is to typically assume n is three and secondly to rely on "studentized" residuals from an initial estimation of all the data to sort cases into the three groups. Having proposed this alternative, our next task is to evaluate it relative to the existing approaches.

MONTE CARLO SIMULATION

The principle behind Monte Carlo simulation is that the behavior of a statistic in random samples can be assessed by the empirical process of actually drawing multiple random samples and *observing* this behavior. The strategy for doing this is to create an artificial "world," or *pseudo-population*, which resembles the real world in all relevant respects (Mooney, 1997). For our research objectives, we are going to create a pseudo-population of extreme behaviors.

General Structure (high, middle, low)

Since we are interested in extreme behavior, we assume that there are many agencies or organizations that have one input and one output. For any one particular agency, if it generates a bigger output for a similar level of input than many other agencies, it will be considered a best agency that represents the positive extreme behavior. On the contrary, an agency with much lower output for similar levels of input will be considered a worse agency. To simplify the problem, we assume there are only three kinds of linear behavioral models shown as Equation 7, 8 and 9, where y is the dependant (output) variable, x is the independent (input) variable, and e is a random noise term. The slope coefficient represents the level of output per unit input thus high performers generate 20 unit of output per input and low performers loose 5 units of output per unit of input.

Best Group $$y_b = 200 + 20\, x_b + e \qquad (7)$$

Moderate Group $$y_m = 120 + 10\, x_m + e \qquad (8)$$

Worst Group $$y_w = 100 - 5x_w + e \qquad (9)$$

Symmetric vs Non-symmetric sub-samples

To evaluate the different capabilities for the three analytical techniques at identifying extreme behavior, the three groups of data are organized into one single dataset that consists of 600 observations. The three independent variables were generated using a uniform distribution defined in the interval 0 to 10. And the three sets of random errors were generated based on a normal distribution with zero mean and variance 5. Under these parameters, the three equations can be run 200 times respectively and then 200 observations are obtained for each of the three underlying groups. This situation generates a uniform distribution of cases in the data. This is the case originally studied by Wu et. al. (2001). In order to consider a broader range of conditions, different ratios of best, middle and worse cases were simulated. Table 1 summarizes eight symmetric the cases and table 2 summarizes five asymmetric cases, for a total of thirteen alternative distributions of best, middle and worse cases.

Table 1: The symmetrical subgroup structure of simulated datasets

Best-Middle-Worst Ratio	Corresponding Observations in the simulated Datasets		
	Best	Middle	Worst
059005	30	540	30
108010	60	480	60
206020	120	360	120
304030	180	240	180
333333	200	200	200
353035	210	180	210
402040	210	120	240
451045	270	60	270

Table 2: The asymmetrical subgroup structure of simulated datasets

Best-Middle-Worst Ratio	Corresponding Observations in the simulated Datasets		
	Best	Middle	Worst
102070	60	120	420
203050	120	180	300
503020	300	180	120
602020	480	120	120
702010	420	140	60

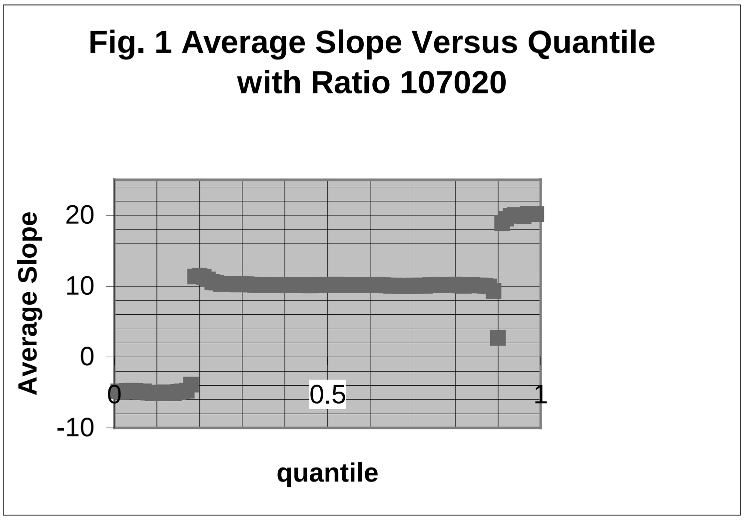
Comparisons

For each of the thirteen cases defined in tables 1 and 2, 100 separate random samples were created for each, thus 1300 separate datasets were generated. For each of the 1300 datasets, two alternative approaches were applies. First, the Quantile Regression models were run on each of the 1300 datasets estimating each conditional quantile from 1 to 99. Next the simplified Segmentation technique was applied assuming three groups and using initial "studentized" residuals to form three groups. Since this approach is heavily dependent on identification of observations into groups in the same way SWLS is, we feel that the use of segmentation captures most of the major characteristics of SWLS especially from an implementation perspective. Thus we did not implement SWLS on these data. The estimation of sub-models for the segmentation approach used standard least squares estimation.

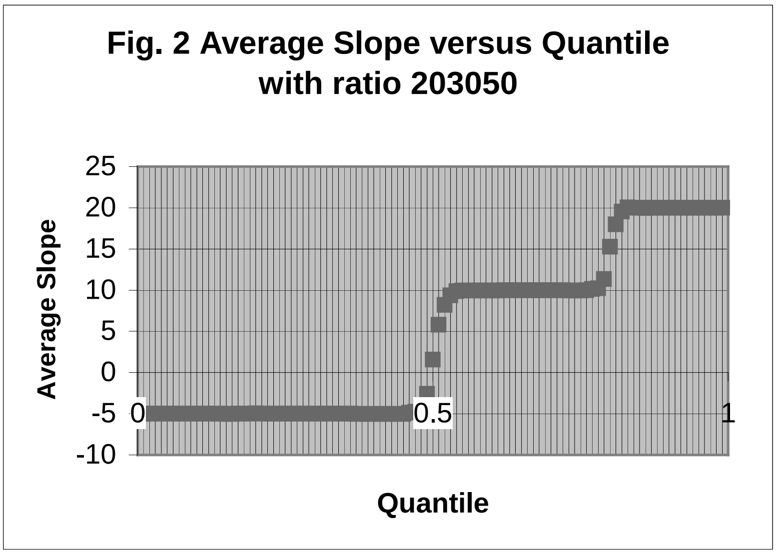
RESULTS & ANALYSIS

Quantile Regression

Prior Monte Carlo results based on symmetric distributions of best, middle and worse for Quantile Regression are reported by Wu et. al. (2001), hence we will focus our attention here on the results from the asymmetric simulations. Figures 1 through 5 present the average estimate of the slope for each of the conditional quantiles where the average is based on the 100 random samples.



The estimation process applied here was able to identify the asymmetric situation in the data. While estimation was problematic at the points in the data were discrete shifts occurred between best and middle or middle and worst, overall quanitle regression was successful, on average, of identifying best, middle and worst cases in the data. Note how in figure 1 the first through eight quantile slopes where at -5 and by the 10th quantile it was up to 10. Note that while the 90th quantile estimate was problematic (e.g. 5) estimates from the 91st through 99th quantile were close to the true value of 20.



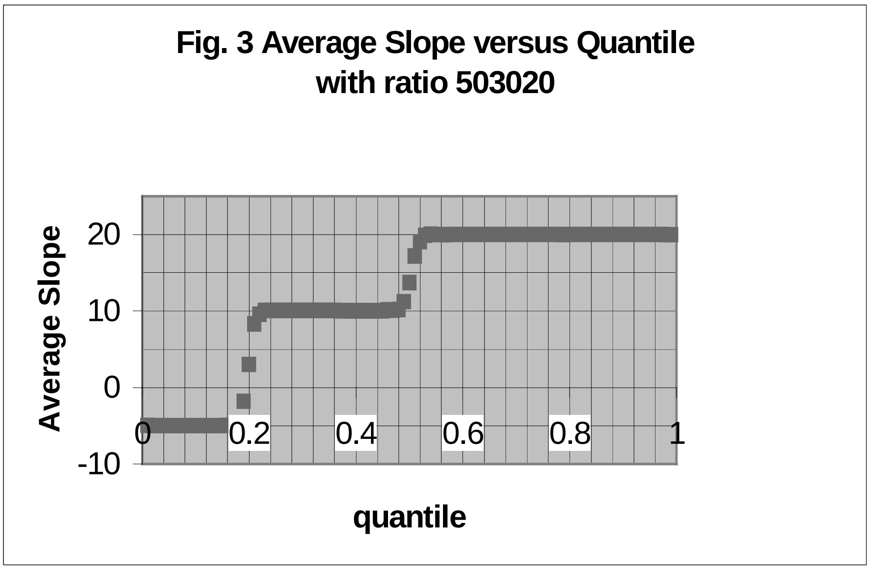


Fig. 3 Average Slope versus Quantile
with ratio 503020
Average Slope
20
10
0
-10
0
0.2
0.4
0.6
0.8
1
quantile

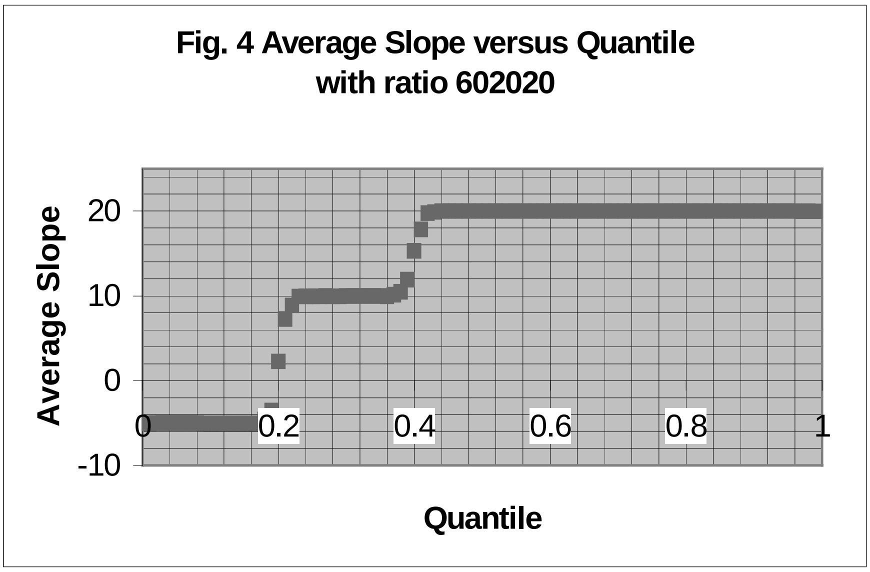


Fig. 4 Average Slope versus Quantile
with ratio 602020
Average Slope
20
10
0
-10
0
0.2
0.4
0.6
0.8
1
Quantile

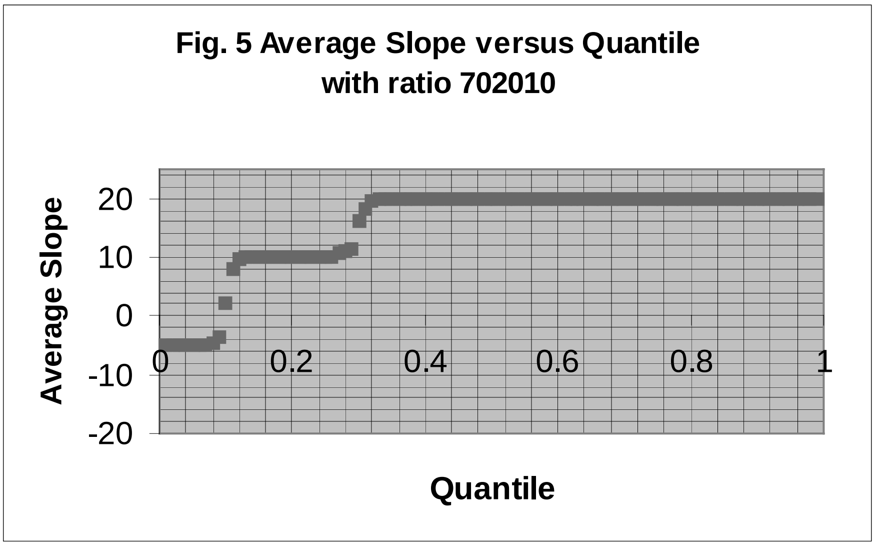


Fig. 5 Average Slope versus Quantile
with ratio 702010
Average Slope
20
10
0
-10
-20
0
0.2
0.4
0.6
0.8
1
Quantile

Segmentation

Out analysis of the segmentation technique includes both the symmetric and asymmetric datasets. The results, shown in Table 3 and Table 4, demonstrate that segmentation works very well when the dataset is symmetric. However, when the simulated dataset are non-symmetric, the segmentation results are significantly worse. Even in the first case where 70% of the observations are in the worst category, the average estimate over the 100 sample is biased (e.g. -3.14 with a standard deviation of .43 when the true value was -5). Segmentation did better for the best cases once 20% or more of the 600 cases were included in the best case category.

Table 3 The Segmentation Outcomes with the structural change of symmetric dataset

BMW Ratio	Best				Middle				Worst			
	Intercept		Slope		Intercept		Slope		Intercept		Slope Numobs	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
059005	199.9	1.67	20.03	0.30	119.84	0.44	10.02	0.08	99.45	1.69	-4.93	0.31
108010	199.9	1.30	20.00	0.22	119.10	0.61	10.13	0.10	99.25	1.44	-4.89	.24
206020	199.9	0.93	20.00	0.15	115.49	1.63	10.64	0.24	100.00	1.42	-4.86	.21
304030	200.6	.84	19.92	.15	114.61	7.09	10.68	1.00	98.90	1.08	-4.85	0.17
333333	200.7	.94	19.91	.15	117.84	9.25	10.16	1.28	98.85	1.07	-4.84	.17
353035	201.0	.82	19.84	.14	120.76	9.30	9.71	1.32	98.96	.96	-4.86	.15
402040	201.1	.85	19.84	.14	131.18	12.2	8.07	1.54	98.62	.99	-4.81	.15
451045	201.3	.97	19.82	.15	140.78	10.8	6.58	1.12	98.32	.99	-4.78	.15

Table 4 The Segmentation Outcomes with the structural change of asymmetrical datasets

BMW Ratio	Best				Middle				Worst			
	Intercept		Slope		Intercept		Slope		Intercept		Slope Numobs	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
102070	219.89	9.33	4.57	1.71	111.19	2.16	-5.23	.78	82.43	4.07	-3.14	.43
203050	201.18	4.92	19.61	1.56	87.17	2.26	13.44	.32	96.45	1.51	-4.56	.21
503020	202.48	1.03	19.68	.15	175.54	8.94	2.77	1.10	93.36	11.82	-3.17	3.43
602020	205.26	1.64	19.35	.22	200.82	9.88	4.57	3.50	78.01	11.23	2.99	3.89
702010	215.55	3.72	18.34	.40	194.49	4.22	20.68	.72	110.3	5.28	5.46	1.30

SUMMARY OF RESULTS AND CONCLUSIONS

Extreme behavior occurs and exists within empirical data, though most of our analytic techniques tend to hide them from view. This suggests that there must be some theoretically meaningful heterogeneity within the associated empirical dataset. Quantile Regression, SWLS and Segmentation can have some roles in isolating and identifying such extreme behavior. However, owing to their different approaches to estimation, and inference, these methods perform this task in different way and with differing abilities..

Best and worst – winners and losers

Quantile Regression, which is median-based regression technique, requires no *apriori* identification of the number of subset or sorting of cases. It provides unique estimates for all conditional quantiles from 0.01 to 0.99. It works equally well when the underlying distribution of high middle and low performers are symmetric or asymmetric. Moreover, the slope estimate from quantile regression is reliable and accurate in the asymmetrical dataset. These demonstrate that quantile regression does have the strong capabilities to locate the extreme behaviors among different datasets.

Segmentation and SWLS similar approaches

We have in this paper suggested a general strategy of segmentation. While this approach is limited by the need to identify in advance the number of segments and sort cases into groups we proposed a simplified heuristic based on the same strategy used by Meier and Gill to solve these same problems. This strategy performed well as long as the underlying distribution of cases between best, middle and worse cases was symmetric. For the non-symmetrical datasets, however, this strategy did not perform well. Some estimates were clearly biased and in general the estimates exhibited much more variation that those generate by Quantile regression.

The possibility of 'Might': the probability of Symmetrical Datasets

'The methods used in public administration and public policy should be determined by the needs of the analyst and the desire to make policy recommendations, not by tradition in writing econometric textbooks.' (Meier and Gill, 2000). While we support this general methodological view, we still must apply rigorous evaluation of any alternative technique suggest. These results suggest that the segmentation approach and possibly SWLS will have some real problems identifying best and worst case in the presence of asymmetric distributions of performance within an empirical dataset. Early work has already suggested that small numbers of cases in the tail of a distribution are problematic for both quantile regression and SWLS. Such results need to be replicated and expanded to develop better implementation strategies to realize Meier and Gill's ultimate goal.

Implementation suggestions of what to use when

On the basis of the above analysis, we produce Table 5 to sum up the comparisons among the three methods. Generally speaking, quantile regression is a reliable analytic tool that can locate the different models, including the best, moderate, and worst. However, segmentation technique, based on least square regression to analyze the subsets, can only generate good estimates with symmetrical data sets.

Table 5: Characteristics of Methodologies for Estimating Extreme Behaviors

Method	Model	Estimation	Inference	Application
Quantile Regression	Natural Structure	Median-based	Better than Mean-based	All kinds of dataset
SWLS	Dichotomized Structure	Mean-based	Does not make sense	All kinds of datasets
Segmentation	Segmented Structure	Mean-based	Mean-based	Symmetrical datasets

Future Directions

- Perform further research on Quantile Regression to investigate the tolerant limit on the smallest ratio of extreme behaviors and the effects of multiple input structures.
- Applying Quantile Regression and Segmentation to practice locating extreme performers.

Jiannan Wu is an Associate Professor and Director of Center for Master of Public Administration at Xi’an Jiaotong University, China: jnw@mail.xjtu.edu.cn.

Stuart Bretschneider is a Professor of Public Administration and Director of the Center for Technology and Information Policy at the Maxwell School of Citizenship and Public Affairs: sibretsc@maxwell.syr.edu.

NOTES

[1] GPRA: the Government Performance and Results Act of 1993 (GPRA, Public law 103-62); NPR: National Performance Review.

[2] GASB: the Governmental Accounting Standards Board (GASB).

[3] According to its authors, SWLS is just one special form of Substantively Weighted Analysis Techniques.

[4] Only in large samples, usually have this relation.

ACKNOWLEDGEMENT

This article is based on a working paper developed at The Maxwell School in January 2001, which was partially funded by a postdoctoral research grant from the Dean of The Maxwell School and the Center for Technology and Information.

REFERENCES

Bassett, G. and Koenker, R. 1978 "Asymptotic Theory of Least Absolute Error Regressions." Journal of the American Statistical Association. 73: 618-622.

Beck, N. 2000. Political methodology: A welcoming discipline, Journal of the American Statistical Association; Alexandria;

Behn, Robert D. 1993 "Case-Analysuis Research and Managerial Effectiveness: Learning How to Lead Organization Up Sand Dunes," in ed. Barry Bozeman, Public Management: The State of the Art. Jossey-Bass. San Francisco.

Blattberg, R. and T. Sargent. 1971. Regression with non-gaussian stable disturbances: Some sampling results. Economerica 39: 501-510. (Cited by Scharf, 1998).

Bloomfield, P., and W. Steiger. 1980. Least absolute deviations curve fitting. SIAM Journal on Scientific and Statitical Computing 1: 290-301.

Bloomfield, P., and W. Steiger. 1983. Least absolute deviations: theory, applications, and algorithms. Boston, Massachusetts: Birkhause Verlag.

Buchinsky, M. 1994. "Changes In The U.S. Wage Structure 1963-1987: Application of Quantile Regression ". Econometrica. 62/2, 405-458.

Buchinsky, M. 1995. "Estimating the Asymptotic Covariance Matrix for Quantile Regression Models; A Monte Carlo Study". Journal of Econometrics. 68/2, 303-338.

Buchinsky, M. 1998. "Recent Advances in Quantile Regression Models". The Journal of Human Resources. XXXIII/1, 88-126.

Buchinsky, M. 1998. "The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach". Journal of Applied_Econometrics. 13, 1-30.

Conley, T.G. and Galenson, D.W. 1994. "Quantile Regression Analysis of Censored Wealth Data." Historical Methods. 27/2, 149-165.

Dielman, T and Pfaffenberger, R. 1982. "LAV (Least Absolute Value) Estimation In Linear Regression: A Review" Journal of Studies in the Management_Sciences. 19, 31-52. (Ask Rick for this article)

Eide, Eric R. and Mark H. Showalter. 1999. Factors affecting the transmission of earnings across generations: a quantile Regression approach, Journal of Human Resources, 34 /2, 253

Fitzenberger, B. 1998. "The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantile Regressions". Journal of Econometrics. 82/2, 235-287.

Friedlander, D. and Robins, P.K. 1997. "The Distributional Impacts of Social Programs" Evaluation Review 21/5, 531-553.

Griffiths, W.E; Hill, R.C; and Judge, G.G. 1993. Learning and Practicing Econometrics. John Wiley & Sons, Inc. New York, NY.

Katorobo, James. 1998. The Study of Best Practices in Civil Service Reforms, Management Development and Governance Division, January: http://magnet.undp.org/docs/psm/best7.htm

Koenker, R. 2000. "Quantile Regression," International Encyclopedia of the Social Science, October/

Koenker, R. and Bassett, G. 1978a. “The Asymptotic Distribution of the Least Absolute Error Estimator”. Journal of the American Statistical Association. 73, 618-622.

Koenker, R. and Bassett, G. 1978b. “Regression Quantiles”. Econometrica. 46: 33-50.

Koenker, R. Galton, Edgeworth, Frisch, 2000. Prospects for Quantile Regression in Econometrics, Journal of Economtrics, 95: 347-374.

Marc-Aurele, Frederick J., Jr., Lynne C. D' Amico, and Stuart Bretschnerider, 2000. Examining the "Extremes" of Bureaucrati Performance: A Comparative Examination of Quantile Regression and Substantially Weighted Analytic Techniques, Working article, Center for Technology and Information Policy, Maxwell School of Citizenship and Public Affairs.

Mata, J. and Machado, J. 1996. “Firm start-up size: A conditional quantile approach”. European Economic Review. 40: 1305-1323.

Meier, K.J. and Keiser, L.R. 1996. “Public Administration as a Science of the Artificial: A Methodology for Prescription.” Public Administration Review 56/5: 459-466.

Meier, Kenneth J.; Wrinkle, Robert D.; and Poliand, J. L., 1999a "Equity and Excellence in Education: A Substantively Reweighted Least Squares Analysis." American Review of Public Administration 29: (Mar.): 5-18.

Neter, John et al, 1996. Applied Linear Statistical Models (Fourth Edition),

Overman, E. Sam and Boyd, Kathy J. 1994, "Best Practice Research and Postbureaucratic Reform." Journal of Public Administration Research and Theory, 4/1: 67-83.

Rocheleau, Bruce. 2000. Prescriptions for Public-sector Information Management: A Review, Analysis, and Critique, American Review of Public Administration, 30/4, December.

Scharf, F.S.; Juanes, F.; and Sutherland, M. 1998. “Inferring Ecological Relationships From The Edges of Scatter Diagrams: Comparison of Regression Techniques.” Ecology. 79/2: 448-460.

Taylor, L. D. 1974. Estimation by minimizing the sum of absolute errors. Pages 169-190 in P. Zarembka, editor, Frontiers in econometrics. Academic Press, New York, New York, USA (cited by Scharf, 1998).

Theobald, Nick and Jeff Gill, 1999. A Tale of Two States: Comparing Educational Performance Between Texas and California Using SWAT Analysis, Presented in the 1999 Annual Meeting of the Midwestern Political Science Association, April 14-17, Chicago.

Western, B. 1995. "Concepts and Suggestions for Robust Regression Analysis." American Journal of Political Science 39: 786-817.

ABOUT IPMR	
IPMR	The International Public Management Review (IPMR) is the electronic journal of the International Public Management Network (IPMN). All work published in IPMR is double blind reviewed according to standard academic journal procedures. The purpose of the International Public Management Review is to publish manuscripts reporting original, creative research in the field of public management. Theoretical, empirical and applied work including case studies of individual nations and governments, and comparative studies are given equal weight for publication consideration.
IPMN	The mission of the International Public Management Network is to provide a forum for sharing ideas, concepts and results of research and practice in the field of public management, and to stimulate critical thinking about alternative approaches to problem solving and decision making in the public sector. IPMN includes over 600 members representing sixty different countries and has a goal of expanding membership to include representatives from as many nations as possible IPMN is a voluntary non-profit network and membership is free.
Websites	IPMR: http://www.ipmr.net/ (download of articles is free of charge) IPMN: http://www.inpuma.net/