RALPH'S PRETTY GOOD GROCERY VERSUS RALPH'S SUPER MARKET:

SEPARATING EXCELLENT AGENCIES FROM THE GOOD ONES

Jeff Gill

California Polytechnic University

and

Kenneth J. Meier

Texas A&M University

Published in Public Administration Review 61 (January/February), 9-17.

A previous version of this article was presented at the 1998 Midwest Political Science Association annual meetings, Palmer House Hotel, Chicago, IL. We would like to thank Stu Bretschneider for comments on an earlier draft. This article is part of a series of studies on SWAT techniques. All data and

documentation to replicate the analysis can be obtained from the authors. Those articles not in print and updates of previous work can be accessed at http://www.calpoly.edu/~jgill.

Charles Goodsell (1983) characterizes American bureaucracy as similar to a good used car; it is reliable, relatively inexpensive to operate and generally gives good service. Bureaucracy, in Goodsell's estimation, is the equivalent of Ralph's Pretty Good Grocery in Lake Woebegone. Goodsell's assessment is meant to be an average for American bureaucracy; he recognized, and others have presented information, that American bureaucracies vary a great deal about that average (Meier 1993; Wolf 1997). With the current public philosophy of neoconservative economics (Lan and Rosenblom, 1991; Osborne and Gaebler 1992), being pretty good is unlikely to be good enough to meet the expectations of policy makers and the public. Public administrators can no longer be content to simply distinguish the good from the bad and the ugly. What is needed is a way to look at the pretty good agencies and distinguish among that group those that are exceptional, the best organizations that others can emulate to provide more effective performance.

Two general theories explain why the performance of bureaucracy varies. A general set of approaches, best characterized as open system theory (Thompson 1967; Rourke 1984), argues that some organizations perform better because they are more highly skilled, possess more useful expertise, or use better quality leadership to exploit their environment to a higher

degree. These agencies are able to do more with less (Downs 1967). Because organizations make decisions by satisficing (Simon 1947), however, even among organizations performing at the highest levels, there will be considerable variation in what they do and how they do it.

A second, more provocative theory of organizational variance and survival is Herbert Kaufman's (1991). Kaufman argues that organizations survive and flourish not because they do things better than other organizations, but rather because they are lucky. In the Darwinist natural selection of organizations, the survivors are those blessed with favorable and stable environments.[1]

Within the context of either theory, being able to distinguish good (lucky) agencies from better or exceptional (really lucky) agencies is well worth while. At the same time if high performance can be attributed primarily to factors in an organization's environment, then Kaufman is likely to be correct

[1]The logical extension of Kaufman's argument is Chubb and Moe's (1989) suggestion that we should focus on organizational environments and structure them in such a way that they do not make a large number of contradictory demands on the organization. In their study of school choice, Chubb and Moe argue that suburban schools and private schools perform better because they exist in homogeneous environments.

about why some organizations out perform others;[2] and our attention should shift to the organization's environment and how to structure this environment. If, on the other hand, high performance results because the best agencies use environment inputs more effectively than other agencies, the open system's perspective will gain credence.

This article, as a result, has both practical and theoretical ends. It uses recently developed substantively weighted analytical techniques (SWAT[3]) to focus on performance optimization by agencies. The units of analysis are 534 school systems in Texas. School systems are not only the most prevalent public bureaucracies in the United States, but they also have some

---

[2]Kaufman could still be wrong if the key organizational skill that is operating is the ability to placate one's environment. As open systems, bureaucracies both respond to their environment and shape the nature of that environment. What might appear to be a highly favorable environment, might actually reflect exceptional political and managerial skills. As one illustration, J. Edgar Hoover was able to define the FBI's role in such a way to be able to deal with cases that were highly visible yet easy to solve (bank robbery, kidnapping) and avoid cases that were difficult to solve or corrupting (prohibition, drug abuse); see Poveda (1990). The result of this strategy was a great deal of organizational autonomy and ample resources.

[3]Software for a variety of platforms can be obtained free of charge from our webpage. See http://www.calpoly.edu/~jgill. We also provide additional documentation on the technical background and various diagnostic techniques that SWAT practitioners may be interested in reading.

relatively objective measures of outputs so that questions of performance can be addressed. First, the article presents a statistical model where student performance is a function of the school system's environment and various policies and resource allocations. Second, after a brief introduction to SWAT techniques, we modify that approach to distinguish or benchmark "excellent" districts from those that are merely good. Our benchmark is a relative standard not an absolute one; we are selecting agencies that do more with less. Third, we examine the agencies with these techniques, both illustrating how the techniques can be used for this purpose and also addressing the theoretical dispute about how organizations change and adapt. Finally, we discuss additional elements of organizational performance that can be assessed in a similar manner.

Modeling Organizational Performance

An education production function is a model that portrays school districts as economic organizations; they receive inputs (resources and students) from their environment and produce outputs (educated students among others). Despite this seemingly straightforward idea, a vast literature has developed an endless variety of education production functions (Burtless 1996; Smith 1995; Hanushek 1986;1989;1996). Because our objective is to contribute to the literature on public organizations rather than

the education policy literature, our discussion of the models and all their possible nuances is brief (for a more elaborate discussion see Burtless 1996). We are seeking to set up a benchmark for evaluation rather than to resolve substantive issues in education policy.

Our outcome variable, that is, the measure of school system outputs, is based on student scores on standardized tests. Texas requires all school districts to administer exams to students on an annual basis. Our outcome variable is the percentage of students who passed these exams in 1991.

The explanatory variables fall into four general types--environmental constraints, financial resources, teacher qualifications, and district policies. Environmental constraints are factors that restrict agency performance; in the case of education the key constraint is how difficult/easy it is to educate students. The three measures of constraint, all correlated with poverty, are the percent of poor students (defined as those eligible for free school lunches), black students, and Latino students. Each should be negatively related to organizational performance.

Financial resources are the basic raw materials of any organization's attempt to meet its goals. Three measures of financial resources are included--per student instructional funds,

average teacher's salary, and percent of funds received via state aid. These represent total resources devoted to education, the attractiveness of teaching positions in a competitive market place, and state efforts to overcome the unequal distribution of local financial resources. All relationships should be positive.

The two teacher qualification measures are the percent of teachers who hold a temporary certification in a subject specialty (as opposed to a permanent certification) and the average number of years of teacher experience. The relationship for noncertification should be negative, but the predicted relationship for experience is ambiguous (see Meier, Gill and Waller 1999).

Finally, the education production function contains three policy measures--the percentage of students taking gifted classes, class size, and student attendance (percent attending on an average day). Performance should be positively related to gifted classes and attendance and negatively related to class size.

Texas has a large number of school districts; many are very small or deal with a homogeneous student body. In an effort to use a set of organizations relatively similar in the tasks that they perform, we have restricted our analysis to the 534 school districts with a least 500 students and between 10 and 90 percent Anglo students.

SWAT Approaches to Analysis

The SWAT approach to analysis begins with the assumption that different organizations vary in their ability to use resources in pursuit of goals. Rather than using regression diagnostics to avoid unusual organizations (i.e., outliers) and generalize to the average case, SWAT exploits the information in the outlying cases to determine how agencies that perform better than average differ from those that are average.

The SWAT approach starts as a statistical procedure (linear regression), but is actually a qualitative investigative technique. The focus is not on parameter estimation or statistical reliability. Instead, SWAT highlights *relative* differences in the way that outlying cases (organizations) use explanatory variables (resources). This distinction is important. Using SWAT to produce exact parameter estimates for the purpose of prediction could lead to unreliable results. Using SWAT to isolate and focus on cases with particularly successful outcomes, however, is a qualitatively informed method. Prescriptive advice can be found by highlighting the use and mix of resources that allow organizations to perform better than expected.

In their original study Meier and Keiser (1996) proposed

using jackknifed residuals[4] (see box: **The Mechanics of SWAT**) of greater than +0.7 (from a regression of organizational performance) to select the high performing cases. The agencies that fail to meet this criterion are gradually downweighted in a series of regressions to reveal how high performing agencies differ from average ones. Gill (1997) relates this selection criterion to the F-distribution and shows that a jackknifed residual of greater than 0.7 will essentially select about 20 percent of the cases for most distributions. In short, it will select the "pretty good" agencies rather than the "super star" agencies.

[PAR Gray Box About Here]

This article applies SWAT techniques in a somewhat different manner. Rather than substantively weighting the above average cases, it will gradually change the definition of an above average agency so that it encompasses fewer and fewer agencies. These smaller and smaller subsets of agencies that perform better and better will be the "super markets" of agency performers.

Findings

---

4 A jackknifed residual is nothing more than a measure of how far above or below the regression line a given data point is. Jackknifed residuals are based on standardized distances from a regression line with the point in question excluded from the calculations made for that line.

The ordinary least squares results from the statistical model for all 534 school districts are found in Table 1. These results should be considered the base regression and serve as a standard for comparison.[5] The general predictions of the model are borne out rather well. Student pass rates are negatively associated with all three environmental factors--low income, black, and Latino students. Financial resources do not fare as well; only teachers' salaries are significantly related to student performance although the other two measures are in the predicted direction. The teacher qualification measures are disappointing; both relationships are negative, but neither meets traditional levels of significance. Policy measures hold up rather well with performance being positively related to attendance and gifted classes and negatively related to class size.

[Table 1 About Here]

The Pretty Good Agencies

The most basic version of SWAT, substantively weighted least squares (SWLS), takes the jackknifed residuals from this equation. It then runs a series of weighted least squares regressions down-weighting those cases that do not exceed 0.7. In this case a

[5]Note: p-values are provided for consistency with the literature in this area. See Gill (1999) for a discussion of the problems associated with p-values and "stars" in social science literatures.

series of 19 additional regressions were run and each iteration down-weighted the "average cases" by 0.05, leaving the high performing cases weighted at 1.0. This iterative process continued until the final regression had respective weights of 0.05 and 1.0.[6]

Table 2 presents the final weighted regression from the SWLS analysis. Because our selection criterion for above average was 0.7, this regression essentially shows how the "pretty good" agencies differ from the average agencies. One good way to compare the results of these two regressions is with a multi-regression barplot (see Figure 1). Several findings jump out from this graphic. First, student performance is now positively and significantly related to instructional monies and percent of state aid. Second, student performance in the above average agencies is also negatively related to noncertified teachers. Third, the positive relationship between gifted classes and student performance disappears. Quite clearly, the relationships for the above average agencies differ from those for the average agency. Figure 1 also provides normal-quantile plots of the residuals from each model. This diagnostic is an indication of whether

---

[6]The analyst can vary these weights by either increasing the size of the increment change and decreasing the number of iterations or by decreasing the size of the incremental change and increasing the number of iterations.

heteroscedasticity has been introduced during the SWAT procedure. The absence of deviance from the line indicates homoscedastic, approximately normally distributed residuals for all models.

[Table 2 and Figure 1 About Here]

A more precise view of the difference between average agencies and the pretty good ones is shown in Table 3 which compares the slopes for the two sets of regressions. In addition to the striking findings in Figure 1, there are a variety of incremental differences between the sets of agencies. All other things being equal, the relationships for the above average agencies are 29% larger for low income students, 5 percent smaller for black students, and a fairly large 40% smaller for Latino students.[7] The relationship for teachers' salaries is 30% smaller in the pretty good districts suggesting that salaries per se are not as important in these districts; at the same time instructional funds (now significant) and state aid (+98%) are far more important.[8] The relationship for teacher certification is 31%

[7]Substantively this means the pretty good agencies are more affected by low income students but less affected by minority students. More directly stated, minority students pass rates are higher in the pretty good districts.

[8]Why might state aid be so important to these school districts? In organizational terms one needs to remember that major state aid for education is a relatively recent phenomenon in Texas. Many school districts are like organizations that have been on severe financial diets for many years. The best of these

larger for the pretty good districts. In terms of policies, the pretty good agencies have a 35% smaller relationship for gifted classes, 17% smaller for attendance and 27% larger for class size.

[Table 3 About Here]

What is really interesting in these differences is that the differences matter to policy makers. The environmental forces are fixed and cannot be changed by managers. The good news, however, is that the pretty good agencies have identifiable resource decisions that not only matter, but differ from the average case. This is where SWAT analysis is the most useful: pretty good agencies should not necessarily heed prescriptive advice from an analysis that focuses on the average agency. In this case the OLS results would indicate that spending limited resources to decrease the number of non-certified teachers is not an effective way to increase exam pass rates. For the pretty good agencies, however, the opposite is true. Class size is another interesting case. In both models class size is important, but for the pretty good agencies the predicted effect is 27% greater. So if allocative decisions are being made at pretty good schools, the advice would be that decreases in class size have a higher payoff. Similarly,

---

organizations have lots of ideas for improvement but lack the money necessary to do so. State aid essentially provides new monies and taps this reservoir of built up reforms.

gifted classes only matter for average districts; those in the above average group are unlikely to make any gains by increasing gifted classes.

The Super Agencies

While the contrast between the average districts and the pretty good agencies is valuable, our interest here is the "best" agencies, those that do a great deal better than even the pretty good agencies. To provide leverage on this problem, we essentially repeated the above analysis but changed the benchmark from 0.7 to 0.8. We continued this approach by increasing the jackknifed residual selection parameter by 0.1 eight additional times until the last SWLS regression procedure used a jackknifed residual of 1.6.

Table 4 shows how this process focused on fewer and fewer agencies that performed better and better. The pretty good regression, as noted above, was not all that selective with some 123 of 534 districts qualifying. The pretty good districts still had a significantly higher mean pass rate (65.0) than did all districts (55.6). The number of districts in the top category continues to drop as the jackknifed residual increases until only 21 agencies, less than four percent of the total, remain in the high performing category.

[Table 4 About Here]

What the Super Agencies Do Differently

Table 5 presents the final SWLS results for the best agencies regression (R > +1.6). The slopes in this regression are compared in relative terms to those for all agencies and those for the pretty good agency regression in Table 6 and in Figure 1. Table 5 suggests that the best agencies are affected by far fewer forces in their environment. Latino students are no longer significant; a striking finding suggesting that in these districts Latino students do about as well on the achievement test as Anglo students.[9] Two other variables that are significant for both all districts and the pretty good regression--attendance and teacher salaries--also drop to insignificance. The very best districts do not appear restricted by either their absenteeism rate or the inability to pay higher salaries (note from Table 7 that teachers salaries do not differ between these two groups). The super district regression also repeats the pattern of the pretty good regression that gifted classes do not matter.

[Tables 5, 6 About Here]

For three variables, the relationships for excellent agencies

[9]In substantive terms this is a remarkable finding suggesting that the super districts have found a way to achieve equity in test scores between Anglos and Latinos. Exactly what the districts are doing needs to be probed via a series of in depth case studies.

look similar to those for the pretty good districts--low income students, black students and instructional funding. In the latter case, both the pretty good and the best regression show a significant relationship, but the all agencies regression does not. What distinguishes the best from the pretty good regression are the other three relationships--class size, teacher certification, and state aid. The best districts get twice the impact of reducing class size as the pretty good districts get. The best districts also get about 54% more from better teacher qualifications and 27% more from increases in state aid. If one were to focus on a single variable that appears to distinguish the best districts from the pretty good ones it would be in what these districts do as they reduce the size of their classes.

Given the cogent summary that the final regressions provide of the difference between excellent and pretty good agencies, one might ask was it worthwhile going through the iterative process? Why not simply jump to the final results (assuming that one would not jump too far and not have any districts remaining)? The iterative process contains a great deal of useful information; this illustrated by Figures 2 and 3. Figure 2 charts the change in relationships for percent Latino students and percent attendance as the set of districts becomes more exclusive (i.e., performs better). For Latino students, there is a gap between all

districts and the pretty good districts. Even though the best districts' slope for Latinos is insignificant and the pretty good slope is significant, in substantive terms they are about the same size. The gains achieved by the best districts, therefore, are already apparent in the pretty good districts. A different pattern is shown for attendance. Pretty good districts do not overcome attendance problems with nearly the skill that the best districts do.

[Figures 2 and 3 About Here]

Figure 3 presents the relationship change graphs for state aid, class size, teacher noncertification, and instructional funding. Instructional funding shows a major jump from all districts (where it is nonsignificant) to the pretty good districts. The best districts really do not get much more out of additional instructional funds than the pretty good districts do, suggesting that the process for doing so is fairly well known to districts with above average talents. Certification and class size form a different pattern with the best districts doing much better than the pretty good ones. The real differences appear at jackknifed residuals of +1.0 and higher suggesting that a fairly high level of skills is needed to maximize return from such resources. Finally, the state aid curve shows what appears to be an eventual diminishing marginal return. The impact of state aid,

while higher among the best agencies than the pretty good ones, actually peaks at a jackknifed residual of 1.3.

Lucky or Good?

A key theoretical question concerning the difference between the super agencies and the pretty good agencies is, are they actually better or just lucky? One view of quality versus luck is to determine if some agencies have more favorable environments than the other agencies. If the super agencies have more favorable inputs, then the argument that they are lucky rather than good gains some credence. If the inputs are relatively equal, then the difference is in what the agencies do with their inputs. Translating inputs into higher levels of outputs requires some skill rather than just luck.[10]

Table 7 compares the means for the explanatory variables for the super agencies and the other agencies. Despite the 19 point difference in pass rates, the means of the explanatory variables are relatively similar. In only two cases are the differences statistically significant at the .05 level--class size and per capita instructional funds. The super agencies have a mean class

[10]We are overstating the case somewhat. That is, an agency could try something and by luck get better performance by some quirk of fate. The true test distinguishing between those agencies that are lucky and those that are good requires a longer term assessment. An agency that out performs its peers with the same inputs year after year cannot be tagged with the term

size of 14.2 (compared to 15.4) and spend $165 more in instructional monies per pupil. These are relatively modest differences and can account for no more than two percentage points of the 19 percentage point difference between the two groups. Kaufman's theory that organizations survive because they are lucky does not appear to hold for these agencies. The super agencies differ from the average agency, not because the super agencies are lucky but rather because they are better at turning their relatively scarce inputs into valued outputs.

[Table 7 About Here]

Other Applications

This study was designed to distinguish between public agencies that were "pretty good" at their job from those that were even better. The process used, a SWAT technique altered to focus on better and better performance, can be used in a wide variety of public policy and public management situations.

First, while the focus here was on the best performers, quite clearly the emphasis could also be on the worst performers (see Meier, Gill and Waller 1999). One interesting situation might be where public agencies are interested in preventing the worst case scenario. In environmental protection, for example, the nature of risk assessment is such that policymakers are concerned with what

---

"lucky."

the most extreme cases could be. If adequate models of environmental quality can be constructed, then SWAT techniques can focus on environmental cases with negative residuals and with those further and further below the regression line.

Second, the process can also be used to study regulatory compliance. In all areas of regulation, some firms comply rather quickly with the law. Others are more reticent and some resist compliance with every resource at their possession. With a measure of compliance, SWAT can be used to construct models of the most resistant to compliance (all other things being equal) and to focus on how those firms make decisions different from the average firm.[11]

Third, additional work needs to focus on when the number of "exceptional" cases becomes too small to provide any meaningful information. Quite clearly the selection criterion can be increased and fewer and fewer programs will qualify. The benefit of having fewer programs is that these are more likely to be the elite programs. The disadvantage is that there might be so few of these programs that their activities and processes cannot be

[11]The method likely has some application in the area of deviant behavior. Because our concerns are with how organizations deal with their environments, we leave these issues for others to address.

generalized (or serve as role models) to other public organizations.

Fourth, SWAT techniques are clearly applicable to a wide variety of private sector activities. In any situation where goals are relatively clear and a production function can be set up, the technique can provide a wealth of useful information. Nothing in the public-private distinction prevents the application of this technique in the private sector.

Conclusion

This article had both methodological and substantive goals. Methodologically, we used SWAT techniques to single out elite agencies. The best agencies differed not only from the average agency but from the above average agency. This article provides some guidance for those seeking to reform government by transferring the techniques from the best agencies to all other agencies. The designation of the best agencies should rely on systematic methods that combine quantitative analysis with qualitative assessments rather using anecdotes.

The approach is useful both for theoretical reasons as presented in this article and practical management purposes. SWAT has the ability to assist managers in focusing on both the best performing agencies and to isolate on those factors that contribute to that performance. At that point some detailed

management analysis can focus on these agencies to determine if what they do can be transferred to other agencies.

Substantively, this article addressed why some agencies perform at higher levels than others. Two theories have very different conclusions about why some agencies succeed and some fail. Open systems' theory holds that agencies succeed because they have better leadership, more skills, more adaptable technology and other factors internal to the organization. Kaufman, on the other hand, feels that success is a function of having a favorable environment. In short, while open systems contends that organizations succeed because they are good, Kaufman would content that success if more a matter of luck.

While a supportive environment is always nice for any organization to have, this study found that the best agencies do more with less; they transform inputs into outputs at a much higher rate. If anything, the best agencies are less restricted by their environments than the average agency. Agencies that succeed may or may not be lucky, but they are clearly good.

PAR Gray Box:

The Mechanics of SWAT

We use the elementary form of SWAT in this analysis: SWLS (Substantively Weighted Least Squares). SWLS is based on simple weighted multivariate linear regression which is run 20 consecutive times on the same data, although users are free to vary this parameter. The first iteration weights all data points equal to one (i.e. unweighted OLS regression), followed by regressions that consecutively down-weight by 0.05 each case whose jackknifed residual is less that 0.70. The i^{th} jackkifed residual(also called an externally studentized residual) is the normal residual weighted inversely proportional to the estimate of the regression standard error *leaving out the* i^{th} *case*. The idea is to measure the influence of one particular case on the OLS calculations.

Virtually any statistical software package can perform weighted least squares and thus SWAT.

After 20 iterations high performing cases will have weights at 1.0 and low performing cases will have weights at 0.05. At this point, variables whose regression coefficients are found to be different from OLS slopes are identified as indicating those factors which have a different effect on higher performing cases. SWAT, therefore, identifies cases that perform well above

expectation (high jackknifed residual) given their allotted resources not just high performers due to resource richness.

Why is this distinction between high performing cases and highly advantaged cases identified by SWAT important? First consider the problem of defining high performing cases without a specific methodology. Clearly highly advantaged cases benefit from more resources (i.e., the corresponding high levels of explanatory variables). To find a high performing case, that is, one doing well *given* a specific mix of levels requires the analyst to look at the corresponding residual. Conversely, a highly disadvantaged case may be performing extremely well relative to similarly affected cases but not relative to advantaged cases. In both scenarios, we are interested in residual outliers with all model specified explanations included.

References

Burtless, Gary. 1996. *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C.: Brookings Institution.

Chubb, John and Terry Moe. 1989. *Politics, Markets and America's Schools.* Washington: The Brookings Institution.

Downs, Anthony. 1967. *Inside Bureaucracy*. Boston: Little, Brown.

Gill, Jeff. 1999. "The Insignificance of Null Hypothesis Significance Testing." *Political Research Quarterly* 52 (September).

Gill, Jeff. 1997. "Generalized Substantively Reweighted Least Squares Regression." Political Methodology WWW Site. http://wizard.ucr.edu/polmeth/working_papers97.

Greene, William. 1993. *Econometric Analysis*. New York: MacMillan.

Goodsell, Charles T. 1983. *The Case for Bureaucracy*. Chatham NJ: Chatham House.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24:1141-77.

Hanushek, Eric A. 1989. "The Impact of Differential Expenditures on School Performance." *Educational Researcher* 23 (4): 45-65.

Hanushek, Eric A. 1996. "School Resources and Student Performance." In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, Gary Burtless, ed. Washington, D.C.: Brookings Institution.

Kaufman, Herbert. 1991. *Time Chance and Organizations: Natural Selection in a Perilous Environment*. Washington: The Brookings Institution.

Lan, Zhiyong and David H. Rosenbloom. 1992. "Public Administration in Transition?" *Public Administration Review* 52 (Nov.\Dec.): 535-7.

Meier, Kenneth J. 1993. *Politics and the Bureaucracy*. Pacific Grove: Brooks/Cole.

Meier, Kenneth J., Jeff Gill and George Waller. 1999. "Optimal Performance vs. Risk Aversion: An Application of Substantively Weighted Least Squares." In *Advancing Public Management: New Developments in Theory, Methods, and Practice*. Jeff Brudney, Larry O'Toole, and Hal Rainey (eds.). Georgetown University Press.

Meier, Kenneth J. and Lael R. Keiser. 1996. "Public Administration as a Science of the Artificial: A Methodology for Prescription." *Public Administration Review* 56: 459-66.

Osborne, David and Ted Gaebler. 1992. *Re-Inventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. Reading, MA: Addison-Wesley.

Poveda, Tony G. 1990. *Lawlessness and Reform: The FBI in Transition*. Pacific Grove CA: Brooks/Cole.

Rourke, Francis. 1984. *Bureaucracy, Politics and Public Policy, 3rd*. Boston: Little Brown.

Simon, Herbert. 1947. *Administrative Behavior*. New York: Free Press.

Smith, Kevin B. 1995. "Policy, Markets, and Bureaucracy: Reexamining School Choice." *Journal of Politics* 56 (May), 475-491.

Thompson, James D. 1967. *Organizations in Action*. New York: McGraw Hill.

Wolf, Patrick 1997. "Why Must We Reinvent the Federal Government? Putting Historical Developmental Claims to the Test." *Journal of Public Administration Research and Theory* 7 (July): 353-89.

Table 1.

The Education Production Function--OLS, All Districts

Dependent Variable = Exam Pass Rate

Explanatory Variable	Value	Std Error	t value	P
Intercept	-35.7455	29.2202	-1.22	0.2218
Environment				
Percent Low Income	-0.2931	0.0374	-7.84	0.0000
Percent Black Student	-0.2307	0.0346	-6.67	0.0000
Percent Latino	-0.1146	0.0262	-4.37	0.0000
Financial				
Instructional Funds	0.3534*	1.7715	0.20	0.8421
Teachers' Salary	1.2444*	0.3568	3.49	0.0005
State Aid Percent	0.0303	0.0246	1.23	0.2184
Policy				
Attendance	0.9187	0.2934	3.13	0.0018
Gifted Classes	0.1985	0.0996	1.99	0.0468
Class Size	-0.9083	0.3370	-2.70	0.0073
Teachers				
Non-certified Percent	-0.1256	0.0699	-1.80	0.0728
Teacher Experience	-0.0006	0.2267	-0.00	0.9978

Residual standard error: 7.153 on 522 degrees of freedom
Multiple R-Squared: 0.59
F-statistic: 69.33 on 11 and 522 degrees of freedom, the p-value is 0
*Explanatory variable in thousands of dollars

Table 2:

The Pretty Good Agencies--SWLS

Dependent Variable = Exam Pass Rate

Explanatory Variable	Value	Std Error	t value	P
Intercept	-8.4915	20.3821	-0.43	0.6700
Environment				
Percent Low Income	-0.3778	0.0323	-11.71	0.0000
Percent Black	-0.2198	0.0287	-7.65	0.0000
Percent Latino	-0.0683	0.0235	-2.91	0.0036
Financial				
Instructional Funds	4.3350*	1.6056	2.70	0.0069
Teachers' Salary	0.8696*	0.3011	2.89	0.0039
State Aid Percent	0.0601	0.0207	2.90	0.0037
Policy				
Attendance	0.7616	0.1913	3.98	0.0001
Gifted Classes	0.1297	0.0943	1.37	0.1692
Class Size	-1.1558	0.2839	-4.07	0.0000
Teachers				
Non-certified	-0.1643	0.0545	-3.02	0.0025
Teacher Experience	0.0598	0.1808	0.33	0.7409

Residual standard error: 3.030 on 522 degrees of freedom
Multiple R-Squared: 0.72
F-statistic: 122.06 on 11 and 522 degrees of freedom, the p-value is 0
*Explanatory variable in thousands of dollars

Table 3.

Average Agencies versus the Pretty Good Ones:

A Comparison of Slopes

Dependent Variable = Exam Pass Rate

Explanatory Variable	All Agencies-OLS	Pretty Good Agencies-SWLS	Ratio
Environment			
Percent Low Income	-0.2931	-0.3778	1.29
Percent Black Student	-0.2307	-0.2198	0.95
Percent Latino	-0.1146	-0.0683	0.60
Financial			
Instructional Funds	0.3530	4.3350	*
Teachers' Salary	1.2444	0.8696	0.70
State Aid Percent	0.0303	0.0601	1.98
Policy			
Attendance	0.9187	0.7616	0.83
Gifted Classes	0.1985	0.1297	0.65
Class Size	-0.9083	-1.5578	1.27
Teachers			
Non-certified Percent	-0.1256	-0.1643	1.31
Teacher Experience	-0.0006	0.0598	**

*The OLS Instructional Funds coefficient is essentially 0 rendering the ratio meaningless.
**Neither coefficient significant.

Table 4. Number of Agencies and Average Pass Rate

Jackknifed Residual	N	Mean	Standard Deviation
0.7	123	65.0	10.1
0.8	112	65.6	9.9
0.9	93	66.4	9.5
1.0	73	67.1	10.0
1.1	59	70.7	9.0
1.2	50	68.3	9.2
1.3	39	69.6	9.7
1.4	34	70.8	9.5
1.5	24	73.9	6.7
1.6	21	73.8	6.8

Table 5.

The Excellent Agencies--SWLS

Dependent Variable = Exam Pass Rate

Explanatory Variable	Value	Std Error	t value	P
Intercept	61.0073	25.9944	2.35	0.0189
Environment				
Percent Low Income	-0.3916	0.0525	-7.45	0.0000
Percent Black Student	-0.2814	0.0479	-5.88	0.0000
Percent Latino	-0.0632	0.0366	-1.73	0.0842
Financial				
Instructional Funds	4.6992*	2.3335	2.01	0.0440
Teachers' Salary	0.4992*	0.4729	1.06	0.2912
State Aid Percent	0.0766	0.0291	2.63	0.0086
Policy				
Attendance	0.2760	0.2423	1.14	0.2545
Gifted Classes	0.1392	0.1444	0.96	0.2939
Class Size	-2.3293	0.4591	-5.07	0.0000
Teachers				
Non-certified Percent	-0.2533	0.0833	-3.04	0.0024
Teacher Experience	0.3409	0.3034	1.12	0.2611

Residual standard error: 2.793 on 522 degrees of freedom
Multiple R-Squared: 0.49
F-statistic: 45.63 on 11 and 522 degrees of freedom, the p-value is 0
*Explanatory variable in thousands of dollars

Table 6.

Comparing the Excellent Agencies to All Others

Explanatory Variable	Dependent Variable = Exam Pass Rate Ratio of Excellent Slope to All Agencies	Pretty Good Slope
Environment		
Percent Low Income	1.34	1.04
Percent Black Student	1.22	1.28
Percent Latino	0.55	0.92
Financial		
Instructional Funds	*	1.08
Teachers' Salary	0.40	0.57
State Aid Percent	2.52	1.27
Policy		
Attendance	0.30	0.36
Gifted Classes	0.70	1.08
Class Size	2.56	2.02
Teachers		
Non-certified Percent	2.02	1.54

*variable not significant in OLS equation

Table 7. Super Agencies and the Also Rans: A Comparison

Explanatory Variable	Mean Values: Super Agencies	Others	t value	P
Environment				
Percent Low Income	39.6	40.6	0.25	0.81
Percent Black	9.3	11.5	0.77	0.44
Percent Latino	28.8	31.2	0.38	0.70
Financial				
Instruction Funds	2411.5	2246.7	1.98	0.0073
Teacher's Salary	25785.1	25878.0	0.24	0.81
Percent State Aid	47.8	48.0	0.05	0.96
Policy				
Attendance	96.0	96.1	0.41	0.69
Gifted Classes	6.4	6.7	0.41	0.68
Class Size	14.2	15.4	3.08	0.0021
Teachers				
Non-certified	5.3	5.2	0.14	0.89
Experience	12.1	11.4	1.83	0.067
Dependent Variable				
Student Pass Rates	73.8	54.8	8.14	0.0000