
Does China's Civil Service System Improve Government Performance?

A Case Study of Education Bureau of Ningbo City

Zhiren Zhou, Peking University
Haitao Yin, University of Pennsylvania
Feng Wang, University of Southern California

Abstract: Performance improvement sits at the heart of the study of public administration. Performance improvement requires performance measurement and relies heavily on effective management of human capital. This paper addresses both performance measurement and management of human capital in the context of China.

China introduced its civil service system in 1993 with performance improvement as its ultimate goal. After years of implementation and practices, we attempt to make an overall assessment of the reform with a straightforward question "Has the civil service system improved government performance?"

Taking the Education Bureau of Ningbo City as a case, our research design begins with efficiency measurement of the bureau and Data Envelopment Analysis (DEA) is applied as the measurement tool. A comparison of agency performance before and after the introduction of the civil service system is carried out to obtain the basic judgment on the effects of the reform. Then seven contributing factors are ranked on the basis of structured focus-group interviews with civil service reform as one of them. It is found that limited efficiency gain was achieved in the Education Bureau and civil service reform made little impact on agency performance. Some theoretical explanations to those findings are provided. We hope that the study not only provides a case for assessment of the civil service reform against its stated goals, but also sheds light on the use of DEA method in efficiency measurement.

Performance improvement sits at the heart of the study of public administration. Wilson (1887) argued that one of the two objects of administrative study is to discover how government can do the proper things with the utmost possible efficiency and at the least possible cost either of money or of energy. In spite of the vigorous advocacy by the school of New Public Administration, ethics never takes the place of efficiency and effectiveness¹ (Yin and Wang, 2000; Lui, 1994). The past two decades witnessed the rapid growth of performance management. In response to a series of economic pressures and social dissatisfactions with public services, governments, especially those in OECD countries, moved away from the traditional focus on inputs and control toward a focus on results and performance.

Performance measurement is a key component of performance management. As Armstrong (1994) has argued, you cannot improve performance if you cannot measure it. However, there are genuine difficulties in measuring performance. An extensive literature has underscored these difficulties and emphasized the dangers of misleading performance indicators (OECD, 1994; Downs and Larkey, 1986; Turner, 1995).

Performance improvement relies heavily on effective management of human capital. Many scholars have explored the relationship between personal management and government performance. For example, Downs and Larkey argued that improving the motivation and skill of the government workforce seem

to be a promising area for productivity gains given the limited opportunities to improve the productivity of agencies through capital substitution and the creation of new work procedures and organizational forms. And this is especially true if one believes that the lassitude of government workers is a chief source of inefficiency (Downs and Larkey, 1986). In their performance management movements, most OECD countries adopted new personal management approaches in the public services which stemmed from their clear recognition that more effective management of people in areas such as pay and employment practices, working methods, organizational culture and job satisfaction will lead to more effective and efficient public service organizations (U.S. Office of Management and Budget, 2002; Maguire and Lidbury, 1996).

This paper addresses both performance measurement and management of human capital in the context of China. China introduced its civil service system in 1993 after years of pilot experiment and debate. The official goals of the reform were set "to optimize the civil service workforce, curb corruption and improve administrative efficiency and effectiveness (*Xiaoneng*)" (State Council, PRC, 1993; Ministry of Personnel, PRC, 1993).

After years of implementation and practices, it is now possible as well as desirable to make an overall assessment of the personnel system reform with a straightforward question "Has the civil service system improved government performance?"

Our research design begins with performance

measurement and evaluation of selected government agencies. A comparison of agency performance before and after the introduction of the civil service system is carried out to obtain the basic judgment on the effects of the reform. Then seven contributing factors are ranked on the basis of structured focus-group interviews with civil service reform as one of them. The ranking of the civil service reform among the seven is thus an illustration, albeit subjective, of its relative importance in agency performance improvement in the eyes of stakeholders. As a case study, this paper focuses on the Education Bureau of Ningbo in Zhejiang Province, a large city with a population of more than 5 million. For reasons discussed in the next section, efficiency is our main concern and Data Envelopment Analysis (DEA) is applied as the measurement tool which allows the use of scattered data from various sources.

The paper proceeds in five sections. Section 2 clarifies the terminology, discusses the methodology and describes the data. Special attention is paid to the DEA tool, including the basic idea and procedures, data requirement and its advantages and limitations. Section 3 reports our main findings. Section 4 offers some theoretical explanations to those findings. And the paper concludes with a summary of our study. It is our hope that the study not only provides a case for assessment of the civil service reform against its stated goals, but also sheds light on the use of DEA method in efficiency measurement.

Terminology

Performance and related concepts such as economy, efficiency, effectiveness and productivity are often discussed by scholars, politicians, government officials and others interested in public affairs, but the meaning of these terminologies are more often assumed than explicitly defined (Prachyapruit, 1994; Lui, 1994). While volatile and imprecise concepts might be acceptable — even desirable — for politicians, it is unimaginable for policy analysts to measure performance unless it is clearly defined. So our first effort is to find way out of the maze of terminologies.

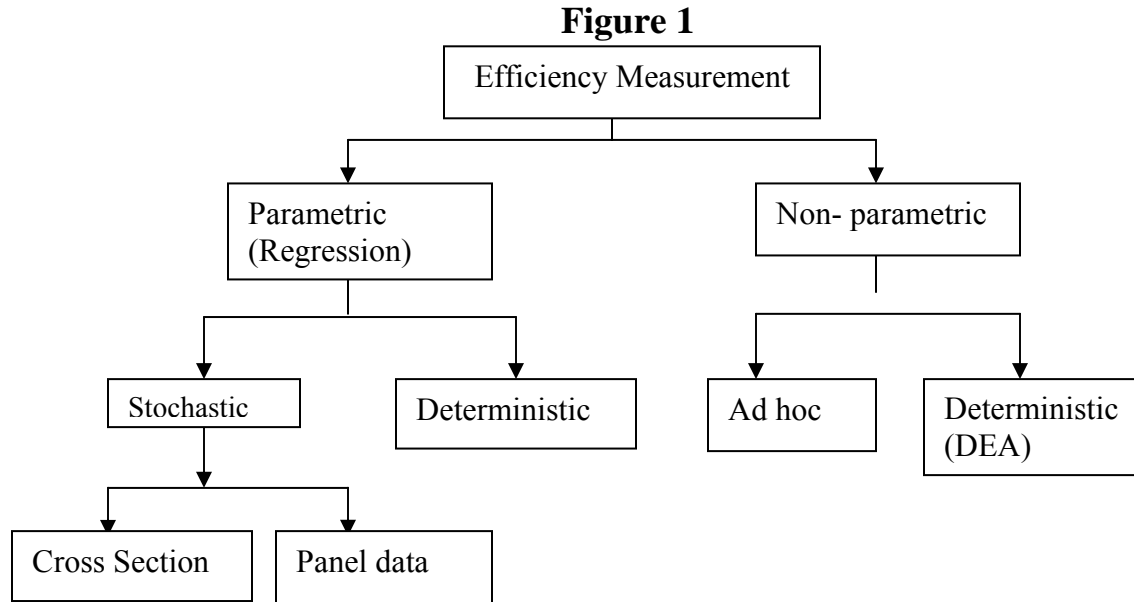
An extensively accepted framework for performance measurement is the “3Es” recommended by the British Treasury, in which performance is decomposed into three indicators: Economy, Efficiency and Effectiveness (Lewis and Jones, 1990; Zhou, 1995). According to British Treasury, economy describes the extent to which the cost of inputs is minimized, efficiency is defined as the relationship of the output of an activity or organization to the associated inputs, and effectiveness refers to the extent to which output contributes to final outcomes. To make a better

understanding, it is necessary to distinguish “output” and “outcome” clearly. Williams (2003) provides the following definition: “output refers to the immediate material effect of government processes,” while “outcomes are the social changes following outputs or processes.”

Effectiveness is arguably the most important one in that, resources, although organized economically and used efficiently, will be largely wasted if goods and services produced do not achieve their intended objectives (outcomes). However, we will focus on efficiency for the following reasons. First, it is very difficult to measure outcomes. For example, the World Bank (1980) defines the outcomes of education as the external effects of output — that is, the ability of people to be socially and economically productive. Secondly, like most other cities in China, Ningbo’s performance management is still at primitive stage compared with those in OECD countries. The absence of strategic performance planning and systematic performance indicators leads to a genuine lack of performance information and data, which in turn makes sound and comprehensive performance measurement impossible. Thirdly, China’s public sector performance management is still characterized by a top-down approach in which the chief executive (the mayor in the case of Ningbo) decides what outcomes to pursue and what outputs to produce and holds agency heads accountable. A review of the performance contract between the mayor and head of the Education Bureau will find that the performance objectives are quite vague and focused on outputs rather than the achievements of outcomes. And finally, our concern is the effects of one specific program — civil service system reform. However, very few outcomes may be attributed unambiguously to particular programs because in most cases there are several factors affecting a particular outcome (OECD, 1994).

Methodology

The parametric approach requires the imposition of a specific functional form (e.g., a regression equation, a production function, etc.) relating the independent variables to the dependent variable(s). The functional form selected also requires specific assumptions about the distribution of the error terms (e.g., independently and identically normally distributed) and many other restrictions, such as factors earning the value of their marginal product (Charnes, Cooper, Lewin and Seiford, 1994). Considering the highly complex input-output relationships of education (Psacharopoulos and Woodhall, 1985), it seems not a good idea to employ parametric approach in this study.



Sources: Barrow 1990.

Figure 1 presents the different methodologies that have been developed to assess efficiency.

Unlike parametric approaches, DEA does not require any assumption about the functional form.² Now let us illustrate input-oriented³ Data Envelopment Analysis with a simple, one-output, two-input example (Coelli, 1996a, 1996b, 1998). Consider a group of five private colleges that teach computing courses to students (the single output) using staff and computers (the two inputs). If we are willing to assume constant returns to scale (CRS)⁴, we can plot the DEA frontier on a two-dimensional diagram. The data for our simple example is listed in **Table 1**. The input/output ratios for this example are plotted in **Figure 1**, along with the piece-wise linear DEA frontier. The DEA frontier

envelops the data points. All points that lie on the frontier are efficient, while all points that lie within the frontier are inefficient. All the efficient points have efficiency value of “1”. The inefficiency of a particular point, for example college 3, is measured along a ray from the origin to that point. The efficiency of college 3 is 0.833. This is the ratio of the distance from the origin (point 0) to point 3’ (on the frontier) over the distance from the origin to point 3. This implies that college 3 should be able to proportionally reduce the consumption of all inputs by 16.7 percent without reducing output. That is, production at the point denoted 3’ in **Figure 2**.

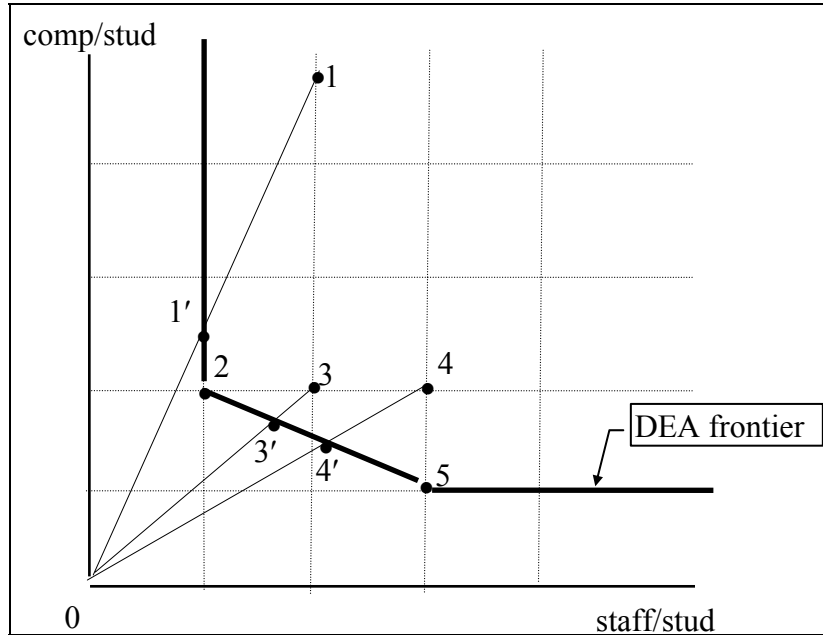
Table 1: Example Data

College	Inputs		Output	Ratios	
	Staff	Computers	Students	Staff/Stud	Comp/Stud
1	2	5	1	2	5
2	2	4	2	1	2
3	6	6	3	2	2
4	3	2	1	3	2
5	6	2	2	3	1

For those interested in DEA methodology, please refer to Coelli and Rao (1995, 1998), Lovell (1993), Ganley and Cubbin (1992), and Charnes, Cooper, Lewin and Seiford (1994). What I want to highlight is that DEA calculation tends to show the governmental agency in a favorable light in the following sense. It produces only

relative efficiency measures, because efficiency scores are generated from *actual* observed data for each Decision Making Unit (DMU). As a result, the efficiency value “1” does not say the DMU is absolutely efficient, but says it is efficient relative to all other DMUs in the observed data.

Figure 2: CRS Input-Orientated DEA Example



The biggest challenges in this study are raised by the small sample size. The first one is how to defend the reliability of our results. Second is how to discern the efficient points (the points on the efficient surface/frontier). If we cannot discern the efficient points, the power to compare different DMUs would be significantly reduced with the small sample, since all the DMUs defining the frontier have the same efficiency value of “1”. To ensure the reliability of our results, we include robustness test by changing the number and types of input and output measures.

We adopted the following methods to address the second challenge. First of all, we employ Constant Returns to Scale (CRS) model, not Variable Returns to Scale (VRS) model. CRS can dramatically decrease the number of DMUs on the efficient surface (frontier).⁵ Second, we rank the efficient points. The earliest attempt to address this problem is due to Andersen and Petersen (1993). Their paper presented a procedure to evaluate an efficient DMU by comparing it with a linear combination of all other DMUs in the sample, i.e., the DMU itself is excluded. The efficiency value attached to the efficient DMU indicates how much it may increase its input vector proportionally while preserving efficiency. (Recall the efficiency value attached to the inefficient DMU indicates how much it has to decrease its input vector proportionally to reach the efficient

frontier). Thanks to a Lovell and Rouse (2002) working paper, Andersen and Petersen’s ideas not only become operationalized, but we are able to use the conventional DEA software to compare the efficient DMUs. The software we used is DEAP version 2.1 provided by the center for Efficiency and Productivity Analysis, the University of New England⁶.

Data

To conduct Data Envelopment Analysis, we need data on inputs and outputs. So the first job is to look for the variables that can best describe inputs and outputs of Education Bureaus. Two considerations dominated our variable selections. First, DEA studies suffer from a problem similar to the “degrees of freedom” problem in statistical analyses. That is, with a small sample size, such as 10 in our study, one can only consider a small number of inputs and outputs. If one included too many variables (i.e., inputs and outputs) into the DEA model, one would observe that the majority of data points would be on the frontier. As a result, we will lose the ability to compare different data points. In this study, only two output variables and two input variables are employed. We also do the DEA in the case of one-output and two inputs to test the robustness of our results.

The second consideration is data availability. It imposes the biggest difficulty for conducting empirical study, especially in developing countries. This difficulty limits our selection of variables. Our quantitative data are mainly from *Ningbo Statistics Yearbook* and *Zhejiang Statistics Yearbook*. Other data are from our 1999 survey in *Education Bureau of Ningbo*, in which face-to-face interviews were conducted.

Finally, we select the following variables to describe inputs and outputs. The number of enrolled students is used to describe the *quantity* of output. How to incorporate information regarding the *quality* of the output is a great challenge to measure efficiency (Harry et al., 1979). The World Bank defined education as achievement of pupils or students, which refers to knowledge, skills, behavior, and attitudes — as measured by tests, examination results, and the like (World Bank, 1980). However, Psacharopoulos and Woodhall argued that such tests may be difficult or expensive to administer or may be regarded as unreliable measures of *quality*. Scholars often measure it in purely quantitative terms, such as the number of graduates or qualified school dropouts produced in the education system.⁷ In our empirical study, we employed the percentage of junior high school students who enter senior high school to describe the quality of output (*graduation rate* hereafter). Since our aim is to measure the efficiency of the Education Bureau (*Jiaoyu Ju*) in Ningbo city, not junior high schools, this is an extremely rough indicator. However, this is the best indicator we can use. First, we do not know how many senior high school graduates entered colleges. Second, almost all the elementary school students can enter the junior high schools since the nine-year compulsory education regulations. So this is not a meaningful indicator. Third, a comprehensive report provided by Education Bureau of Ningbo (*Jiaoyu Ju*) views this variable as the best indicator for its performance.

Two input variables used in our empirical study are the number of school employees (including teachers and administrative staff) and the education expenditure. To test the robustness of our study, we also employ the number of schoolteachers as one input indicator. A little trouble arises in collecting education expenditure data. For 1993-1994 and 1996-2000, we can get it directly from the *Ningbo Statistics Yearbook* and *Zhejiang Statistics Yearbook*. However, we only have the expenditure data combining education, culture, and sanitation for 1991, and expenditure data combining education and culture for 1992 and 1995. We adjusted these data in the following ways. We obtained the

percentage of education and culture expenditure in the expenditure combining education, culture and sanitation from 1992 data (percentage 1), and the percentage of education expenditure in the expenditure combining education and culture from 1993 data (percentage 2). Then we multiply the 1991 data by percentage 1 and percentage 2 to get the education expenditure in 1991. We multiply 1992 data by percentage 2 to get the education expenditure in 1991.⁸ In the same vein, we adjust the 1995 data, multiplying it by the average percentage of education expenditure in the total expenditure combining education and culture of 1994 and 1996.

Results

Table 1 illustrates the results obtained from the Data Envelopment Analysis based on input-oriented CRS model. Four cases are designed to test the robustness of our results. Case 1 is the basic model, in which school employees and education expenditures are employed as input variables, and enrolled students and graduation rates are used as output variables. Case 2-4 are the variants of case 1. In case 2, we use schoolteachers as input variable instead of school employees, that is, we exclude the administrative staff in schools. In case 3, we only use enrolled students as an output indicator. And in case 4, we only use graduation rates as output. In **Table 1**, we report the results from both non-ranking efficient points procedure and ranking efficient points procedure, which are abbreviated into NREP and REP respectively. The inefficient points have the same efficiency score under the two procedures, and the efficient points with score “1” under NREP were ranked by the method Lovell and Rouse recommended under REP.⁹

Ningbo City introduced its new civil service system in 1995. In order to examine the effects of that introduction on government efficiency, we need to look at whether there exists significant difference between efficiency scores before 1995 inclusively and those after 1995 exclusively. T-tests are conducted, and **Table 2** reports the mean, standard deviation, t-value and p-value. Obviously, there is no significant difference between the efficiency scores before 1995 inclusively and those after 1995 exclusively. This conclusion is robust across different cases and procedures. Now we can answer the question raised by the title, that is, in terms of efficiency, the introduction of China's civil service reform has no significant influence on governmental performance. At least this is true for the Education Bureau (*Jiaoyu Ju*), Ningbo City.¹⁰

Table 1: Results from DEA

	Case 1		Case 2		Case 3		Case 4	
	NREP	REP	NREP	REP	NREP	REP	NREP	REP
1991	1.000	1.256	1.000	1.256	1.000	1.016	1.000	1.256
1992	1.000	1.136	1.000	1.096	1.000	1.136	0.969	0.969
1993	0.972	0.972	0.993	0.993	0.972	0.972	0.885	0.885
1994	0.951	0.951	0.982	0.982	0.951	0.951	0.877	0.877
1995	0.943	0.943	0.958	0.958	0.938	0.938	0.903	0.903
1996	0.964	0.964	0.984	0.984	0.906	0.906	0.957	0.957
1997	0.972	0.972	0.980	0.980	0.946	0.946	0.948	0.948
1998	1.000	1.040	1.000	1.048	1.000	1.040	0.919	0.919
1999	1.000	1.056	1.000	1.048	0.955	0.955	1.000	1.040
2000	0.983	0.983	0.982	0.982	0.852	0.852	0.983	0.983

Table 2: Results from T-test

	Case 1		Case 2		Case 3		Case 4	
	NREP	REP	NREP	REP	NREP	REP	NREP	REP
Mean Difference	0.0106	-0.0486	0.0026	-0.0486	-0.0404	-0.0628	0.0346	-0.0086
Std.Error	0.0140	0.0649	0.009	0.0574	0.0279	0.0474	0.0282	0.0742
T-value	0.76	-0.75	0.29	-0.85	-1.4470	-1.3253	1.2277	-0.1159
P-value*	0.470	0.475	0.781	0.422	0.1859	0.2217	0.2545	0.9106

* Two side t-test for data before 1995 inclusively and data after 1995 exclusively.

Table 3

Factors/Order	Important (1-3)	Neutral (4-5)	Unimportant (6-8)
<i>Increase of Education Expenditure</i>	3	1	3
<i>Technical Progress</i>	2	1	4
<i>Management by Objectives</i>	3	3	1
<i>Political Support from above</i>	6	0	1
<i>Introduction of Civil Service System</i>	1	4	2
<i>Perfection of Laws and Regulations</i>	3	2	2
<i>Management Improvement</i>	1	3	3
<i>Public Support</i>	2	0	5

Note: The numbers in the cells indicate the number of the interviewees who view the corresponding factor “important,” “neutral,” or “unimportant.” For example, the number “3” in the upper-left corner cell indicates that three respondents think the factor “Increase of Education expenditure” important in influencing government performance.

Source: Yin, 1999.

This conclusion is also confirmed by our field survey conducted in 1999. Seven officials in the Education Bureau (*Jiaoyu Ju*) were interviewed based on a structural questionnaire. The answers to question,

“Please rank the following factors in terms of their contribution to improve government performance,” are summarized in **Table 3**.

Of the seven interviewees, only one listed

“introduction of civil service system” as an “important” factor affecting government performance. This survey also shows that “political support and attention” has the most powerful influence. Among the seven interviewees, six listed it as an “important” factor affecting government performance. In the same field survey, a similar question also was asked to five clients.

Table 4 reports the answers. Again, only one of them listed “establishment of civil service system” as an important factor affecting government performance, while four of them view it as “unimportant.” Again, “political support and attention” seems to have the most powerful influence. Four of five interviewees list it as an important factor and only one thinks it is unimportant.

Table 4

Factors/Order	Important (1-3)	Neutral (4)	Unimportant (5-7)
<i>Increase of Education expenditure</i>	3	1	1
<i>Technical Progress</i>	0	1	4
<i>Political Support from above</i>	4		1
<i>Establishment of Civil Service System</i>	1		4
<i>Perfection of Laws and Regulations</i>	3	1	1
<i>Management Improvement</i>	2	2	1
<i>Public Support</i>	2		3

Note: The numbers in the cells indicate the number of the interviewees who view the corresponding factor “important”, “neutral” or “unimportant”. For example, the number “3” in the upper-left corner cell indicates that three respondents think the factor “Increase of Education expenditure” important in influencing government performance.

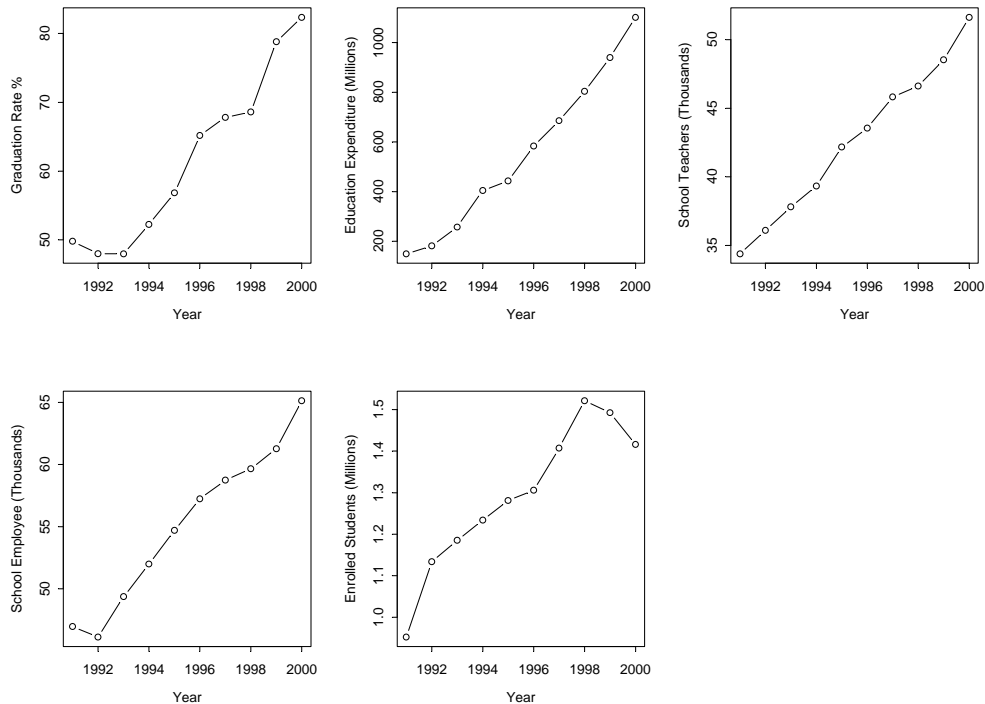
Source: Yin, 1999.

Let’s return to **Table 1**. Combined with **Figure 2**, some comments run as follows: 1) As mentioned before, a comprehensive report by Education Bureau (*Jiaoyu Ju*) of Ningbo views the percentage of junior high school students who enter senior high school as the best indicator for its performance. From **Figure 2**, we can see this indicator increases steadily over years. Case 4 which employed only the graduation rate as output indicator, we also found a rough increasing trend in government efficiency. (It also has the relatively small *p*-value). This suggests that the internal objective of a government agency might influence its behavior and performance in a significant way. A more detailed analysis calls for careful examination of programs initiated by the Education Bureau of Ningbo. However, this is beyond the scope of this paper, and we leave it for future occasions.

2) Case 1–2 outputs combine enrolled students and graduation rate, the lowest efficiency score arises in 1995. A possible explanation is that the agency is changing its output priority from quantity to quality. The increase of enrolled student has slowed down, while the graduation rate still increased slowly.

3) As indicated in **Figure 2**, education expenditure, schoolteachers and school employees have increased smoothly over the years. They do not reflect the changes of the number of enrolled students. This suggests the government budget basically is an outcome of historical conjecture, not an outcome of demand analysis. This is why the government efficiency scores fluctuate most and exhibit a rough decreasing trend in Case 3, which has the number of enrolled students as the only output indicator.

Figure 2: Time Trend of Key Variables



Civil Service Reform and Efficiency Improvement

China’s civil service system reforms initiated systemic adjustments in the rules pertaining to job definition or classification, deployment, job security and membership, reward structures and wage rules. Theoretically, these changes should have a major impact on government employees and further on organizational efficiency and effectiveness (Wise, 1996, as quoted from Burns, 2001). As indicated in the introduction, the increases in efficiency and effectiveness are also what the reformers expected. However, our empirical studies are not positive in this regard. That is, the efficiency of the Education Bureau (*Jiaoyu Ju*) in Ningbo City is not improved significantly after the introduction of civil service systems. Why did this happen? Here are some explanations we are turning to:

Peters (1994, as quoted from Farazmand, 2002) raised three perspectives on administrative reform and reorganization — purposive (top-down) models, environmental (bottom-up) models, and institutional models. Admittedly, China’s civil service reform is a purposive (top-down) one. As Burns (2001) had noted, the impetus for recent administrative reform in China has come at the intervention of paramount leader Deng Xiaoping and the need to propel China to market economy. The biggest challenge for purposive reforms is allowing them to stay only on paper. This is the very problem of China’s civil service reform. Actually, we

can borrow Tocqueville’s comments on the French Revolution to describe China’s civil service reform: the changes it led to are far less than what people imagined, especially in local governments (Tocqueville, 1856; Gilbert [*English version*], 1955).

Most regulations of the new civil service system are the continuance of the old cadre-management regime. And the innovations are implemented in practice mainly by words not by spirit.

The civil service reforms still do not include such concepts as political neutrality and protection from political interference. The party/state forms a single, integrated authority hierarchy to which civil servants are politically responsive. This is why the interviewees (both officials and clients) view “political support” as the most powerful factor influencing governmental performance. Aside from this, few changes can be identified even in technical facets. Entry through open, competitive examinations is viewed as an important advance of the civil service system superior to the old cadre-management regime. However, Ningbo City only organized these examinations twice in 1996 and 1997. In 1997, only five college-students were employed by Education Bureau, and they worked at low-level clerical and technical positions. Cadres in old systems must take qualification examinations to be “transformed” into civil servants. However, these examinations are no more than “nominal” in the sense that only 37 among 19,029

cadres failed and were not allowed to “transformed” into civil servants.

Performance appraisal is the critical component of an incentive mechanism in the new civil service systems. However, it also makes little difference in substance from the old cadre-evaluation regime. The procedure almost stays unchanged (Yin 1999; Yin, 2000). Appraisals comprise only ratings of “excellent,” “satisfactory,” and “unsatisfactory” (Article 25 in Provisional regulations). In practice, less than 1 percent of appraisees receive “unsatisfactory” ratings in the nationwide (Burns, 2001). In Ningbo City, only 0.1 percent and 0.14 percent of appraisees are rated as “unsatisfactory” in 1997 and 1998 respectively (Yin, 2000). According to the internal guidelines, no more than 15 percent of appraisees may receive “excellent” ratings. As a result, most appraisees are doomed to be in the “satisfactory” category, and this significantly reduces the incentive power of performance appraisal.

Provisional regulations mandated that “excellent” ratings should be connected with salary increase and position promotion; however, these connections are extremely weak in practice. One working unit often gives “excellent” ratings to its employees in turns to keep comfortable personal relationships (Yin, 2000). As to compensation,¹¹ the new civil systems reaffirmed “pay according to work” as guiding principles of the civil service salary system (Article 26 in Provisional regulations). As Burns noticed (2001), civil service salaries in China are graded according to levels of responsibility and complexity. This principle has been in operation in China *for decades*. The “new” principles included both rewards for seniority or time-in-service (annual increments) and pay for performance (appraisals-based bonuses). However, because of the nature of the performance appraisal system, the pay-for-performance element is severely attenuated. The appraisals-based bonuses failed to motivate civil servants. After describing the procedure for performance appraisal, one official in Education Bureau (*Jiaoyu Ju*) of Ningbo City concluded that the performance appraisal and the appraisal-based bonus system have no any influence on the organization efficiency (Yin, 2000).

In a nutshell, the introduction of civil service system in Ningbo City makes few changes in personnel management. As a result, we should not be surprised at the conclusion that no significant efficiency improvement is observed. Another lesson we learn is that civil service system reform should be integrated and coordinated with specific organization objectives and other reforms. It is clear from our interviews that civil service reform in Ningbo City only followed the general national reforms and was not linked closely with specific public service objective. In the eyes of Education Bureau officials, it is an ad hoc, top-down reform and has no relationship with their own business. As a result, we cannot expect the hoped-for efficiency gains which relate to the outputs of specific public

service. OECD recommended that high priority should be given to finding ways of integrating human-resource management with the core business of public service (OECD, 1996). We also note that civil service reforms should coordinate with other reforms. After examining the civil service reform in developing countries, Das (1998) argued that civil service reform is more likely to succeed if taken up in the context of structural adjustment. China’s civil service reform fails to integrate itself with broader administrative reforms. In 1998, China’s government began its new administrative reform aiming to streamline and downsize the state. As a result, the open and competitive examination for selecting new civil servants was stopped in Ningbo City.

Furthermore, as a response to financial incentives aiming to encourage government employee to leave, and thus downsize government bureaus, many young and high-educated civil servants go back to university or “drop into the sea” of commerce, which negatively affects the government efficiency.

Now two explanations are suggested for the failure of civil service reforms to improve government efficiency. First, few critical and real changes were made by the purposive (top-down) civil service reforms in Ningbo City. Second, civil service reform fails to be integrated with specific organization objectives and other concurrent administrative reforms.

Data Envelopment Analysis and Efficiency Measurement

Data Envelopment Analysis is a powerful tool to measure efficiency. It requires specific assumptions neither on the functional form relating independent to dependent variable(s) or the distribution of the error terms. Furthermore, since DEA assesses efficiencies along a ray from the origin to the observed production point, the efficiency measures are *units invariant*. That is, changing the units of measurement (e.g. measuring the education expenditure in Million per year instead of Billion per year) will not change the efficiency values. The parametric approaches, however, are not invariant to the measurement units. These properties simplify the data analysis greatly and make DEA attractive.

Our empirical study is limited by the small sample. Otherwise, we can do more. 1) If we have more data, we can do our analysis under Variable Returns to Scale (VRS) model and Non-Increasing Returns to Scale (NIRS) model. Employing VRS model, we can identify scale efficiency. Employing NIRS model, we can further indicate whether the DMU is operating in an area of increasing or the decreasing returns to scale. Based on these analyses, we can give some specific suggestion on operation scale with the objective to improve efficiency. 2) If we have panel data, we may use DEA-like liner programs and a (input- or output-based) Malmquist TFP index to measure productivity change, and to decompose this productivity change into technical change (identifying the natural trend) and technical-efficiency change. 3) If we have more data or

panel data, we can construct some econometric model. Basically, these models will take efficiency or productivity scores obtained from DEA as dependent variable, and the variables affecting efficiency or productivity as independent variables. The advantage of econometric models is that we can identify the factor influencing efficiency or productivity, and raise suggestions based on the results. The purpose of these discussions is not to show off DEA, but to outline some prospects for our project. We leave all of these for future occasions.

The potential challenges to our empirical study not only come from the small sample, but also comes from the nature of the data. Data Envelopment Analysis is first developed to do cross-section data analysis. In our empirical study we use time-series data and treated each DMU as if it were a different DMU for each period. However, the theoretical implications of representing each DMU as if it were a different DMU for each period is a “blind point” in the research of DEA and remains to be worked out. We discussed this problem with Professor Lovell at University of Georgia, who has contributed a lot to the development of DEA. He insisted that there should no problem. However, we still expect that DEA experts pay attention to this problem and give a detailed theoretical explanation. This is important because historical comparison is the basic method to evaluate performance changes in public sector (OECD, 1994).

In all, Data Envelopment Analysis is a powerful and amenable method to assess efficiency. We recommend getting it involved with performance management as a practical technique. Two points worthy of notice are: First, DEA measures efficiency and says little to effectiveness. To evaluate performance more comprehensively, we need to find ways to measure effectiveness. This is a more challenging task. And we have to admit that it might be misleading to focus on improving efficiency while ignoring the difficult task of measuring and improving effectiveness. For example, if you employ graduation rates and enrolled students as output indicators and use the resultant efficiency assessments to evaluate Education Bureaus, no Education Bureau (*Jiaoyu Ju*) will care about developing students’ ability to apply knowledge, let alone building their good virtue, which are more important for a healthy society. Second, we must use caution when we use DEA to assess government efficiency. The details of every case should be examined before conclusions are drawn because different efficiency scores employ different measures of outputs and inputs. The more accurate measures you choose, the more you can say. The selection of outputs and inputs needs to combine the efforts of education experts, schoolteachers, education officials and so on.

Conclusion

In this study, we use Data Envelopment Analysis to measure the efficiency of Education Bureau of Ningbo City. No significant difference was found between the efficiency levels before 1995 inclusively and those after 1995 exclusively. This means that in terms of efficiency, the introduction of China’s civil service reform has no significant influence on governmental performance.

Theoretically, we found that the claimed positive effects of the introduction of civil service systems on governmental efficiency were not achieved in practice. We suggest two explanations for this: First, few critical and real changes were made by the civil service reforms in Ningbo City. Second, civil service reform fails to be integrated with specific organization objectives and other concurrent administrative reforms.

Methodologically, we identify the data requirements, the advantages and limitations of applying DEA. We recommend getting it involved in performance management as a practical technique. Furthermore, we put forward the technical problems DEA experts are expected to address. Finally, we agree that for the sake of a more comprehensive performance evaluation, more efforts should be devoted to assess its effectiveness.

Notes

¹As to the relevant history of productivity measurement in government organizations in USA, please see Downs and Larkey, 1986.

²Ad Hoc methods do not require any assumption about the functional form either. These methods are the oldest tools to measure efficiency, including such approaches as the use of pupil-teacher ratios, pupil-educational budget ratios and so on. The shortcomings of these methods are 1) it is essentially hard to draw any definite conclusions based on ad hoc measures of efficiency. For example, inefficiency could occur when the pupil-teacher ratio were either too high or too low (Barrow, 1990); 2) they do not work in the situation of multi-inputs or multi-outputs.

³There are two approaches to define technical efficiency—one is the input-oriented approach and the other is output-oriented approach. Input-oriented approach addresses the question: “By how much can input quantities be proportionally reduced without changing the output quantities produced?” And output-oriented approach answers the question: “By how much can output quantities be proportionally expanded without altering the input quantities used?” Generally one should select an orientation according to which quantities (inputs or outputs) the decision making units’ managers have most control over. (Coelli, 1996b) In our study, we choose input-oriented approach based on the reasonable assumption that city government and education agencies have more control over education expenditure and school employee than the number of enrolled students. Actually, in many instances you will

observe that the choice of orientation will have only minor influences upon the obtained efficiency scores (Coelli, 1996b).

⁴Alternatively, you could choose Variable Returns to Scale (VRS) model and Non-increasing Returns to Scale (NIRS) model. For the difference between them, please refer to Coelli (1996a, 1996b, 1998).

⁵The price is that we cannot tell the “pure” technical efficiency from the scale efficiency. Since what concern us is the aggregate efficiency assessment of Education Bureau, Ningbo City, a comprehensive technical efficiency is enough for our research purpose. Again for more on VRS, please refer to Coelli (1996a, 1996b, 1998).

⁶<http://www.une.edu.au/febl/EconStud/emet/deap.htm>.

⁷In their paper, Meier, Polinard and Wrinkle used the percentage of students who pass the TAAS test in each school district to measure student performance (Meier, Polinard and Wrinkle, 2000).

⁸Why do we only use 1992 and 1993 data, instead of some averages, to adjust the expenditure data of 1991 and 1992 respectively? The reason is that both percentage 1 and 2 are strictly increasing over years.

⁹All the raw results are available on request.

¹⁰This conclusion is consistent with the World Bank. After examining the civil service reforms in developing countries, Das concluded that the fiscal and efficiency impact was negative; the efficiency gains from the reform in the larger context are not visible (Das, 1998).

¹¹ Pay/incentive systems play an important role in the conceptual framework to guide the future civil services that affect performance (Das, 1998).

Authors' Notes

We are grateful to Professor Lovell at University of Georgia and Professor Coelli at University of New England. They answered our question regarding DEA methods and gave us many helpful suggestions. We are also grateful to two anonymous referees for their valuable comments. Finally, we thank Jiameng Zhang for her help in the data analysis.

Authors

Zhiren Zhou is a professor and vice dean in charge of postgraduate programs at the School of Government, Peking University, China. He is also an adjunct professor of the China National School of Administration and a vice chairman of the National Steering Committee for Public Administration Education. He obtained his Doctoral Degree in public administration from Leeds University, United Kingdom, in 1993. His main teaching and research areas include administrative reform and government innovation, public sector performance management, and public sector economics.

Mr. Haitao Yin is a Ph.D. candidate at the Wharton School, University of Pennsylvania. His current research interests include regulatory economics, public finance and risk management. He earned his MA degree from the School of Government, Peking University and has published 10 papers in various journals in the field of public policy and public administration.

Ms. Feng Wang is a Ph.D. candidate at the School of Policy, Planning and Development, University of Southern California. Her current research focuses on public-private partnerships, public organization and management and environmental management in China. She earned her MA degree from the School of Government, Peking University and has published three papers in various journals in the field of public administration.

Reference:

Andersen, Per and Niels C. Petersen (1993), 'A Procedure for Ranking Efficient Units in Data Envelopment Analysis,' in *Management Science* Vol.39, No.10, October 1993.

Armstrong, Michael (1994), *Performance Management*, London: Sage.

Barrow, Michael M., (1990) 'Techniques of Efficiency Measurement in the Public Sector,' in Cave, Martin, Maurice Kogan and Robert Smith (eds.) (1990), *Output and Performance Measurement in Government: The State of the Art*, London: Jessica Kingsley Publishers.

Burns, John P. (2001) 'The Civil Service System of China: The Impact of the Environment,' in Burns, John P. (ed.) (2001), *Civil Service Systems in Asia*, Northampton: Edward Elgar Publishing Limited.

Charnes, Abraham, William Cooper, Arie Y. Lewin and Lawrence M. Seiford. (1994), *Data Envelopment Analysis: Theory, Methodology, and Application*, Boston: Kluwer Academic Publishers.

Coelli, Tim, D. S. Prasada Rao and George E. Battese (1998), *An Introduction to Efficiency and Productivity Analysis*, Boston: Kluwer Academic Publishers.

Coelli, Tim (1996a), 'Assessing the Performance of Australian Universities Using Data Envelopment Analysis,' *Working paper*.

Coelli, Tim (1996b), 'A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program,' *CEPA Working Paper No. 8/96, ISBN 1 86389 4969*.

Das, S.K. (1998), *Civil service reform and structural adjustment*, New Delhi; New York: Oxford University Press.

- Downs, George W. and Patrick D. Larkey. (1986), *The Search for Government Efficiency: from Hubris to Helplessness*, Philadelphia, PA: Temple University Press.
- Farazmand, Ali (2002), *Administrative reform in developing nations*, Westport, Conn: Praeger.
- Hennessey, Peter (1990), 'The Political and Administrative Background,' in Cave, Martin, Maurice Kogan and Robert Smith (eds.), *Output and Performance Measurement in Government: The State of the Art*, London: Jessica Kingsley Publishers.
- Harry P. Hatry, Sumner Clarren, Therese Houten *et al.* (1979), *Efficiency Measurement for Local Government Services—Some Initial Suggestions*, Washington, D.C.: Urban Institute.
- Henderson, Stewart (1990), 'Performance Measurement and Review in Local Government' in Cave, Martin, Maurice Kogan and Robert Smith (eds.) (1990), *Output and Performance Measurement in Government: The State of the Art*, London: Jessica Kingsley Publishers.
- Irwin, Tim (1996), 'An Analysis of New Zealand's New System of Public-Sector Management,' in OECD (1996), *Public Management Occasional Papers No. 9: Performance Management in Government: Contemporary Illustrations*, Paris: OECD.
- Lewis, Sue and Jeff Jones (1990), 'The Use of Output and Performance Measures in Government Departments,' in Cave, Martin, Maurice Kogan and Robert Smith (eds.), *Output and Performance Measurement in Government: The State of the Art*, London: Jessica Kingsley Publishers.
- Lovell, Knox, Lawrence C. Walters and Lisa L. Wood (1994), 'Stratified Models of Education Production Using Modified DEA and Regression Analysis,' in Charnes, Abraham, William Cooper, Arie Y. Lewin and Lawrence M. Seiford. (eds.) (1994), *Data Envelopment Analysis: Theory, Methodology, and Application*, Boston: Kluwer Academic Publishers.
- Lovell, Knox and A. P. B. Rouse (2002), Equivalent Standard DEA Models to provide Super-Efficiency Scores, *Working Paper*.
- Lui, T. (1994), 'Efficiency as a Political Concept in the Hong Kong Government: Issues and Problems,' in Burns, John P. (ed.) (1994), *Asian Civil Service Systems: Improving Efficiency and Productivity*, Singapore: Time Academic Press.
- Ma, Stephen K. (1996), *Administrative Reform in Post-Mao China: Efficiency or Ethics*, Lanham, MD: University Press of America.
- Maguire, Maria and Christine Lidbury (1996), *Integrating People Management into Public Service Reform*, Paris: Organization for Economic Co-operation and Development.
- Meier, Kenneth J., Polinard, J.L., and Wrinkle, Robert (2000), 'Bureaucracy and Organizational Performance: Causality Arguments about Public Schools', in *American Journal of Political Science*, Volume 44, Issue 3, July 2000.
- Ministry of Personnel, PRC (1993), *Guanyu Yinfa "Guojia Gongwuyuan Zhidu Shishi Fang'an" de Tongzhi* (Circular on "Implementation Plan for Provisional Regulations on National Civil Service System").
- Organization for Economic Co-operation and Development (OECD) (1994), *Public Management Occasional Papers No. 3: Performance Management in Government: Performance Measurement and Results-Oriented Management*, Paris: OECD.
- Psacharopoulos, George and Woodhall, Maureen (1985), *Education for Development: An Analysis of Investment Choices*. New York: Oxford University Press.
- Prachyapruit, Tin (1994), 'The Productivity of Thailand's Elite Civil Servants: Prolegomenon on Problems and Problem Resolutions,' in Burns, John P. (ed.) (1994) *Asian Civil Service Systems: Improving Efficiency and Productivity*, Singapore: Time Academic Press.
- State Council of People's Republic of China (1993), *Guojia Gongwuyuan Zanxing Tiaoli* (Provisional Regulation on National Civil Service System).
- Turner, Michael (1995), *Performance Measurement of Municipal Services: How are America's Cities Measuring Up?* Philadelphia, PA: Pennsylvania Economy League, Eastern Division.
- U.S. Office of Management and Budget (2002), *The President's Management Agenda*, available at <http://www.whitehouse.gov/omb/budget/fy2002/mgmt.pdf>.
- Williams, Daniel W. (2003), 'Measuring Government in the Early 20th Century,' in *Public Administration Review*, November/December 2003, Vol. 63, No. 6.
- Wilson, Woodrow (1887), 'The Study of Administration,' in *Political Science Quarterly* Volume II, No. 2, June 1887.

Yin, Haitao (1999), 'Gongwuyuan Kaohe Zhidu De Neshenghua: BeiJing Beitaipingzhuang Jiedao Banshichu De Ge'an Yanjiu (The Internalization Process of Performance Appraisal of Civil Service: a Case Study of Beitaipingzhuang Sub-district Office, Beijing)', *working paper*.

Yin, Haitao and Feng Wang (2000), 'Lun Xingzheng Liuli Zai Xingzhengxue Yanjiu Zhong De Qiushi (Why is Ethics Undervalued in Public Administration: a Theoretical Elaboration)', in *Journal of Yunnan Administration College*, Vol.5, 2000.

Yin, Haitao (2000), 'Dui Ningboshi Jiaoyuju De Anli Yanjiu (A Case Study on Education Bureau in Ningbo City)', *Working Paper*.

Zhou, Zhiren and Liqun Xie (1995): 'Xingzheng Xiaolv Ceding De Jishu Yu Fangfa (Measuring Administrative Efficiency: Methods and Techniques)', in *Journal of Peking University (Special Issue on Political Science and Public Administration)*, 1995.

Zhou, Zhiren (2000), 'Gonggong Bumen Zhiliang Guanli: Xinshiji De Xinqushi (Quality Management in the Public Sector: New Trends in the New Century)', in *Journal of National School of Administration*, No. 2, 2000.

Zhou, Zhiren (2000), 'Xingzheng Xiaolv Yanjiu De Sange Fazhan Qushi (Three Trends of Development in the Study of Administrative Efficiency)', in *Chinese Public Administration*, No. 1, 2000.