Adoption or implementation? Performance measurement in the City of Guangzhou’s Department of Education

Meili Niu

Introduction

Over the past few decades, performance budgeting has become a widely used management instrument to improve budgeting accountability in both developed and developing countries (Economic Commission for Africa 2003; ADB 2006; Hatry 2006; OECD 2007; Bouckaert and Halligan 2008; Wescott et al. 2009; Niu and Ho 2014). China is no exception to this trend.

In 2003, Guangdong province launched six pilot programs to measure the performance of government programs after they had been fully implemented (Niu et al. 2006). Since then, many Chinese local governments—including provinces, prefectures, counties and districts—have measured the performance of government programs.
Performance budgeting reform in China is a unique budgeting innovation and has been used now for more than 10 years. This makes it one of the longest lasting budgetary reforms in China, even though it was debatable whether China was ready for performance-based budgeting (PBB) (Ma 2005; Wu and Niu 2010).

Despite its wide application globally, the PBB process in developed countries varies considerably and is always challenging, especially when it comes to improving budget decision-making (de Lancer Julnes and Holzer 2001; Andrews 2005; Curristine 2005; Shah and Shen 2007; Wanna 2010; Ho and Im 2015; Lu et al. 2015).

Performance measurement is a crucial technique to promote PBB (Martin 1997). Based on the extent to which performance information is used for budget allocations, Shah and Shen (2007) classified PBB into four types: performance-reported budgeting (PRB), performance-informed budgeting (PIB), performance-based budgeting (PBB) and performance-determined budgeting (PDB). Very few countries have adopted PBB or PDB.

Apart from technical difficulties, organisational, cultural and political contexts also make performance evaluation a very challenging reform in the public sector (Shah and Shen 2007; Bouckaert and Halligan 2008; Ho and Im 2015). Therefore, an examination of PBB requires a deep understanding of the governance context (Niu and Ho 2014; Ho and Im 2015). De Lancer Julnes and Holzer (2001) find that rational and political factors have different impacts on the utilisation of performance measurement.

Inspired by Cronbach et al. (1981) and Beyer and Trice (1982), de Lancer Julnes and Holzer (2001) argue that, to understand the use of performance measurement in the public sector, there is a need to separate the two stages of utilisation into an adoption stage and an implementation stage.

Adoption refers to ‘the development of measures of outputs, outcomes, and efficiency’ and implementation represents ‘the actual use of performance measures for strategic planning, resources allocation, program management, monitoring, evaluation, and reporting to internal management, elected officials, and citizens or the media’ (de Lancer Julnes and Holzer 2001: 695).
This distinction is particularly useful when researching China’s PBB reforms. Due to the fragmentation of the Chinese budgeting system, the primary concern in budget accountability is controlling the line agency’s spending behaviour (Ma and Niu 2007). Since the late 1990s, when China launched its departmental budget reform (DBR), line agencies have been the focus of budgetary reform. Making line agencies more accountable is the key to improving the accountability of the whole Chinese budgeting system.

PBB reform in China is also fragmented. Accountability in many developed countries, such as the United States, Australia and New Zealand, can be achieved through legislative measures that promote PBB and related financial management reforms. But China’s PBB reform followed a bottom-up process. Finance departments at the local government level took the lead, investing a considerable amount of resources in developing performance indicators and designing the evaluation procedures through administrative orders (Niu 2012).

The bottom-up approach does have limitations. As the central budgeting office, the finance department has the authority to review departmental budget requests and monitor budget implementation, but it does not have any formal power to make line agencies accountable for following the administrative orders it issues. It is an open question whether a line agency merely adopts performance measures or actually uses the performance information to improve its departmental budget management.

Despite this lack of formal power, most of the existing literature on China’s PBB uses the finance department as the unit of analysis. Scholars concentrate on the finance department’s strategies and instruments and their impacts on budget management. What is not examined is how the line agency perceives PBB and actually uses performance measurement.

Current performance measurement, as observed by the reporting finance departments, relies on self-evaluation reports (SERs) prepared by line agencies. If the report is well written, the agency’s performance is usually given high marks regardless of how the program actually performs. A better-performing program that is represented by a poorly written SER may not be given a high score.

To understand how a line agency actually uses performance measurement, more detailed case analysis is required, particularly to determine the line agency’s adoption or implementation level.
This is a single case study. The author selected the education department of the City of Guangzhou for several reasons:

1. Although Guangzhou is not the first city to use PBB, it is one of the more successful cases in China, and its education department is one of the best experimental cases in Guangzhou.

2. While the education department’s budget is not representative of line agencies in terms of its size, it has a number of unique characteristics that demonstrate the impact of technical, organisational, cultural and political factors on budget management.

   • First, China’s education system has a long tradition of performance management; from preschools to universities, teachers are paid, in part, according to performance. For example, student testing results, academic publications and so on are used to measure teacher performance. This makes the concept of performance evaluation easier to accept at the program level.

   • Second, the education department has a large budget compared with most other line agencies, and its financial office must deal with a variety of programs, making it a challenge to evaluate them all.

   • Finally, education services in China are always in the public eye and receive extensive media coverage, making them a political priority. For example, when Guangzhou published all 114 line agencies’ budgets in 2009, the education department’s budget received the most attention from the media. This puts more pressure on the department to improve its services.

3. The author has been collaborating with the education department since 2009 on training staff, designing performance procedures and developing indicators. As a third party, the author has also created a research team to evaluate the performance of a preschool program in 2012. Over the past few years, the author has also worked as an expert at the request of the city’s finance department and the city’s congress to examine the education department’s program management. Therefore, the author has access to program information and is able to observe the evolution of the education department’s efforts to install performance management into its budgeting system.

In the next section, this chapter describes China’s PBB reform, followed by a description of the education department in Guangzhou and why it uses performance measurement. The chapter then explores the extent to which the department has adopted and/or implemented performance
measurement and budgeting, including how its approaches have evolved over time. The final section sets out some conclusions about the key factors affecting the utilisation of performance measurement by line agencies.

**PBB reform in China**

PBB reform in China is regarded as a continuing innovation, following DBR. DBR was launched in the late 1990s as a way to improve line agencies’ budget management through a more transparent and rule-based process.

DBR required line agencies to share information on employees, facilities and programs with the finance department so its oversight would be based on complete information (Niu 2010). Before the DBR reform, the finance department did not have timely access to such information.

DBR also enhanced budget control over line agencies by separating departmental spending into three components: employee payments, operational expenses and program expenditure (Niu 2010). This division is aligned with the way China’s government divides areas of responsibility:

1. **Employee payments** in the public sector are fully regulated by the human resources department. Both the size of a line agency’s workforce and each position within it (and its pay level) are decided by the Commission of Public Sector Reform, so there is no discretion over the funds.

2. **Operational expenses** are difficult to estimate because they include many miscellaneous items. To control total spending, DBR assigns each agency an operational quota per capita. For example, for an agency with 200 employees, if the operational quota is RMB30,000 (A$5,700), the operational budget will be RMB6 million (A$1.1 million) (= 30,000 x 200).

3. **Program spending** is the largest component of a department’s budget. DBR required line agencies to rank the priority of programs based on their importance; however, it did not develop a particular approach to priority setting. Performance—the commonly used concept in a public budgeting system—was not the major concern in China during DBR. This resulted in legitimacy problems for program budgeting.
One problem for program budgeting was that program information was not transparent. The finance department did not have complete information with which to examine an agency’s budget requests. Another problem was that most program budget proposals were very simple and lacked details about cost and performance. Cost–benefit analysis was also absent in the departmental budgeting system. Difficulties in program management also led to problems with regard to resource allocation. For example, many approved programs were unable to proceed due to poor planning, so a large amount of allocated money went unspent (Ma and Yu 2012).

To solve the problems related to program budgets, performance measurement was introduced into the Chinese local budgeting system in the early 2000s. As with many other budgetary reforms, this was initiated by the finance department. As Niu (2012) explained, although there were differences in the application of performance measurement by local governments, some similar instruments were used:

1. To avoid an increased workload and to provide credibility, finance departments contracted third parties to evaluate their programs. These outside contractors were usually university research teams, accounting firms or consulting companies, hired to evaluate program performance.

2. Because PBB involves very complicated, technical and political concerns, PBB reform in China started with post-program evaluations. The purpose of the reform was primarily to improve program implementation rather than allocation efficiency.

3. As mentioned above, line agencies have no discretion when it comes to employee payments and operational spending. Therefore, measurement focuses on programs only. Neither employee payments nor operational expenses are evaluated. Usually, programs with a budget up to RMB5 million (A$950,000) at the provincial level and up to RMB500,000 (A$95,000) at the city or county level were selected for evaluation.

4. Although the indicators used for evaluation are quite different, four dimensions are commonly used: inputs, implementation, outputs and results.

5. Because performance measurement was a new concept for the Chinese public budgeting system, both finance departments and line agencies received intensive training in how to develop performance indicators and write SERs.
After years of experiments at the local government level, beginning in 2011, the Ministry of Finance (MoF) passed several administrative orders to promote PBB in both central and local governments, such as the ‘Interim Approaches on Evaluating the Performance of Fiscal Expenditures’ (2011) and ‘Guidelines on Advancing Performance Budgeting Management’ (2011).

To put more pressure on central agencies and provincial governments, the MoF promulgated the ‘Methods on Assessing Performance Budgeting Management’ (2011), which create an index for the assessment of PBB reforms. Within this index, only 30 per cent of the total score is about program evaluation. The remaining 70 per cent is given to the implementation of the reform. Implementation includes whether the governor has given a public speech on PBB reform, whether the province established clear rules for using PBB, the percentage of the budget applying PBB, staff training on PBB, and so on.

In 2011, the MoF also issued the ‘Work Plan on Performance Budgeting Management’ (2012–15) to set out the medium-term goals of the reform. Because the development of performance indicators has been a daunting challenge, the MoF issued the ‘Framework on Common Indicators of Budgeting Performance Evaluation’ in 2013. The framework was not new to those localities that had used performance measurement for many years, but most local governments had not started PBB by 2013. Nevertheless, issuing this framework became the signal that PBB was the major reform the MoF wanted to promote.

In 2014, China amended its Budget Law. As the highest level of legislation on public budgeting, the amendment requires that both line agencies and each level of government consider the previous year’s performance when making budgeting decisions.

---

1 According to the formal rules, in the Chinese departmental budget system, employee and operational expenses are together called basic spending, which is supposed to maintain the organisation’s daily operations. All other expenses are program spending for achieving specific policy goals. However, in practice, because the quota for operational expenses is usually not sufficient for funding daily operations, the finance department allows line agencies to include some operational expenses in the program budget under ‘operational programs’ (Niu 2010).

In China, the executive arm of government usually initiates public sector reform, with the legislature (people’s congresses) having a much more limited role. Therefore, the tradition is to use administrative orders instead of legislation to launch reform and establish its rules.
Even though performance measurement has become a quite popular practice over the past decade in China, generally speaking, the evaluation results seem to have had very limited impact on budgeting allocation (Niu 2012)—for two reasons. First, line agencies are reluctant to use the instrument because it takes a lot of effort to evaluate programs, and there is almost no connection between the performance result and the next year’s departmental budget. Second, the development of performance indicators is very complicated and information on program implementation is not completely transparent in China. Therefore, both line agencies and finance departments are very cautious about relying on the performance result when making budget allocation decisions (Niu 2012).

To understand whether China has fully implemented PBB or merely adopted PBB mechanisms, and the factors that lead to that choice, the key is to examine what the line agency actually does.

The purpose of PBB reform, as the continuing budgetary innovation after DBR, is to improve program planning and implementation to control line agencies’ budgeting behaviour. For post-program evaluation, a line agency’s SER is the fundamental document used for performance measurement, regardless of which third party conducts the evaluation.

As for pre-program evaluations, the finance department still relies heavily on the information provided by the line agency to review the program budget request. Therefore, how line agencies respond to the reform and how they utilise performance measurement tools will shed the most light on whether PBB is merely adopted (and largely ignored) or fully implemented in the Chinese local budgeting system.

The education department of Guangzhou

The education department is one of the 105 line agencies in the City of Guangzhou. In 2016, there were 65 subordinate units within the education department (see Table 7.1) and 25,108 employees, including 10,133 retirees (40.36 per cent of total employees).²

² In China, public sector retirees’ pensions and social benefits are not paid through a unified social security system, but are still part of the line agency’s departmental budget.
Table 7.1 Subordinate units of the education department (departmental budget 2015)

<table>
<thead>
<tr>
<th>Types</th>
<th>Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head administration office</td>
<td>1</td>
</tr>
<tr>
<td>Elementary schools</td>
<td>1</td>
</tr>
<tr>
<td>Preschools</td>
<td>12</td>
</tr>
<tr>
<td>Middle and high schools</td>
<td>11</td>
</tr>
<tr>
<td>Special schools</td>
<td>3</td>
</tr>
<tr>
<td>Vocational high schools</td>
<td>2</td>
</tr>
<tr>
<td>Vocational middle schools</td>
<td>13</td>
</tr>
<tr>
<td>Vocational colleges</td>
<td>4</td>
</tr>
<tr>
<td>Universities</td>
<td>3</td>
</tr>
<tr>
<td>University-affiliated hospitals</td>
<td>7</td>
</tr>
<tr>
<td>Others</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>65</strong></td>
</tr>
</tbody>
</table>

Source: Bureau of Education of Guangzhou Municipality (various years).

Figure 7.1 shows that the education department’s budget was balanced in most years and both revenue and spending grew quickly over the past seven years. In 2016, total revenue was RMB1.4 million (A$264,000), with 98.62 per cent growth compared with 2010 (RMB696.61 million (A$133 million)). Total spending was RMB1.4 billion (A$267 million), with 100.91 per cent growth compared with 2010. And basic spending (on employees and operational expenses) increased from RMB532 million (A$101 million) in 2010 to RMB1.1 billion (A$209 million) in 2016, an increase of 106.66 per cent.

The proportion of program spending decreased from 38.54 per cent in 2012 to 21.5 per cent in 2016. This was primarily due to the growth in the number of retirees and performance pay to employees (Niu and Song 2015a). Another major reason for the smaller proportion devoted to program spending is the reduction in basic construction projects. Guangzhou, as a municipal government, has built many educational facilities and its schools are well equipped. However, the one-child policy has had a shrinking effect on the number of schoolchildren using those facilities and the demand for new infrastructure projects is now low.
Why did the education department use performance measurement?

Guangzhou’s finance department required the education department to use performance measurement and it has done so since 2005, when Guangzhou launched its pilot program.

In 2004, the finance department established its Performance Evaluation Office (PEO) and, in 2005, it issued the ‘Methods for Public Spending Performance Evaluation of Guangzhou (Trial)’. In the same year, the PEO selected two pilot programs: ‘sparrow school’ upgrading and construction of water and road infrastructure for rural areas.

In 2006, the filing of SERs was required for a total of RMB17.5 billion (A$3.3 billion) in program funds across 48 line agencies. The finance department evaluated program performance based on those reports. In 2007, all 400 programs with a budget up to RMB5 million (A$953,000) were evaluated.

---

3 ‘Sparrow schools’ are schools that are relatively small and have poor facilities and low-quality instruction. From 2000 to 2004, Guangzhou invested RMB1.1 billion (A$209 million) to upgrade these schools.
A leadership change in the education department greatly contributed to the use of performance evaluation. From 2005 to 2008, the department was not very interested in developing its own instruments to evaluate program performance. That changed in 2008 when the Ministry of Education in Beijing appointed one of its senior directors, who happened to have a PhD, as the deputy head of the Guangzhou education department. At the same time, the department also appointed to its finance office a key person who happened to have a Masters of Public Administration.

Both these leaders favoured the idea of performance evaluation to improve program management. The deputy head sold the idea to the departmental head and other deputy heads to win the support of the whole leadership team. He also attended all meetings related to performance measurement. This was very important to build the confidence of the financial office and reduce conflicts with subordinate organisations.

External pressure also played a crucial role. Because education services are always in the public eye and education is one of the largest areas of public spending in China, both the media and the legislature like to bring up issues related to education programs.

In 2008, the City Congress of Guangzhou began examining select departmental budgets during its annual plenary session. The education department was one of the first to be examined. Subsequently, it has been one of the departments selected by the congress's deliberation seminar almost every year. Besides the congress, the Mayor of Guangzhou also required more education spending, which further increased the external pressure on budget allocation and program management.

Adoption versus implementation

De Lancer Julnes and Holzer (2001) argue that, to better conceive the use of performance measurement, the two stages of utilisation—adoption and implementation—should be examined separately. Adoption is influenced more heavily by rational factors, while implementation is more heavily influenced by political factors (de Lancer Julnes and Holzer 2001). This distinction helps in the development of a framework to assess the education department's experiment. The two stages are not necessarily part of an evolutionary process. Instead, even though an organisation enters the implementation stage, it could step back to the adoption stage if the context changes (de Lancer Julnes and Holzer 2001).
The adoption stage: 2005–08

In 2005, as one of two departments selected by the finance department for the pilot evaluation, the education department did not develop any specific approaches for its adoption of performance measurement. Instead, it just followed the requirements listed in the ‘Methods for Public Spending Performance Evaluation of Guangzhou (Trial)’ to begin its performance measurement. These included:

1. Establishing a self-evaluation work group, chaired by the departmental head. The secretary of the Commission for Discipline Inspection and the deputy head responsible for financial affairs served as vice-chairmen. The other 13 group members were all senior managers of different offices within the education department.

2. Forming a self-evaluation work plan for the department and submitting it to the city’s Performance Evaluation Work Group4 as a record.

3. Collecting, justifying and analysing the data to evaluate the programs.

4. Calculating the evaluation score and preparing the SER.

One unique feature of China’s performance measurement is that it uses not only the outputs and outcomes but also the input and program management to assess program performance. Until 2014, Guangzhou had been evaluating inputs, program management, outputs and outcomes of program spending. And, because the major reason for Guangzhou using performance evaluation was to reduce the proportion of the budget that remained unspent, the rate of use of budget funds was one of the major performance indicators.

In 2006 and 2007, over 60 per cent of the indicators were input measures. The weights given to input measures were about 30 to 40 per cent at the beginning of the reform. Output and outcome indicators made up about 30 per cent of the total number of indicators, and the remaining 30–40 per cent were program management indicators. Efficiency measures, such as per capita cost, were never indicators of program performance.

---

4 The Performance Evaluation Work Group was established in 2006 to enhance the authority of performance measurement. As well as the finance department, the audit department, the statistics department and the Commission of Discipline Inspection also serve in the group.
Input indicators are financial measures that include the disbursement rate as well as the appropriation arrival rate, the investment multiplier effect, and so on. Program management indicators were developed based on the following procedures: program application and approval, implementation, budget adjustment, program operation and fulfilment. However, due to the diversity of public programs, program management varies across the different types of programs. The finance department required more detailed indicators based on three types of programs: basic construction/renovation, facility purchases and others. Following the finance department’s guidelines, the education department did not develop any other specific approaches to improve performance measurement.

SERs are very important documents for a line agency to demonstrate the results of its programs and the challenges of program operation. In 2006, the finance department required line agencies to submit SERs, including information on line agencies’ responsibilities and organisation, strategic planning, performance goals in the annual work plan, performance achievements, annual program budgets, actual expenses and financial management. Along with the SER, the line agency also had to submit all supporting documents, such as files identifying the establishment of the Performance Evaluation Work Group, the completion document, the audit report for basic construction projects and regulations on financial and program management.

In terms of the completion of the required document and the timeliness of its submission, the education department has always been one of the best-performing departments. One reason for this is that schools usually keep better-quality program records than other public agencies, so it is easier to collect the program data.

However, as with all other agencies, the education department, in the SER section dealing with the challenges for program implementation, stated that shortage of money was the big issue. There was, however, no detailed analysis of why the budget was inadequate and how much more would be needed.

5 In China, it was very common for the line agency or the program to receive less than its full appropriation during the budget year. This was true even though the budget was approved by the legislature. In fact, the finance department is responsible for transferring the money to the line agency’s account and this indicator is meant to evaluate the finance department rather than the line agency. Therefore, it was later removed in some programs.
The implementation stage: 2009–12

The 2009 fiscal year was a watershed for the education department’s performance measurement. Due to the leadership change and the external pressure for transparency and effectiveness of education spending, the department decided to use performance measurement to improve its budgeting management, instead of just fulfilling the finance department’s requirements.

First, the education department extended the scope of its program evaluation. In addition to the programs demanded by the finance department, it also evaluated the performance of selected programs of the subordinate organisations not requiring evaluation by the finance department. And, for those programs evaluated by the finance department, the education department first conducted internal assessments, by inviting experts to comment on its SERs and program operation. The procedures for the two types of evaluations are described in Figure 7.2.

![Figure 7.2 Evaluation procedure for the programs assessed by the finance department](image)

Notes: FD refers to the finance department and ED refers to the education department.
Source: Based on interviews with civil servants at the Education Bureau of Guangzhou.
Figure 7.3 Evaluation procedure for selected programs assessed by the education department

Source: Based on interviews with civil servants at the Education Bureau of Guangzhou.

Figure 7.2 shows that the education department put a lot of effort into improving its program evaluations. This raises the question of whether the external experts were just window dressing for the post-program evaluation or whether they were there to actually help improve the budgeting and program management.

Bischoff and Blaeschke (2015) argued that window dressing is a common reason for imperfect performance management. China is no exception in this. Because performance evaluation relies heavily on SERs and because program information is still not transparent, as long as the SERs are written well, the performance results will look good, whether they are positive or negative. Line agencies invest a lot of time in polishing SERs, and the education department used experts for this purpose as well.

However, the evidence strongly suggests that the education department’s efforts were much more than window dressing. First, the experts also met with program managers, financial officers and school managers and discussed how to improve program and budgeting management.

Second, the education department used experts from universities and the finance department to train staff every year. The training was separated into two components. The first component was about developing performance indicators. Because the programs were classified into three categories—basic construction and renovation, facility purchases and others—the training was separated into three lectures. The second component of training was how to write SERs. This included topics such as how to cite government documents as supporting evidence for a program, how to explain incomplete budget disbursement and how to analyse problems.
related to program operation or financial management. The training was a signal to its subordinate units that the education department wished to use performance measurement tools as substantive instruments to improve budgeting management. The training also helped staff and school managers to understand the implications of performance measurement and improve their organisational capacity for performance management.

Third, to strengthen the impact of performance measurement on budget allocation, the education department had a trial pre-program evaluation and introduced competitive allocation and citizen participation into budget decision-making.

In 2012, the education department selected the Private Affordable Preschools (PAP) program as a pilot for pre-program evaluation. In Guangzhou, the city has been investing in private schools to improve the quality of education services and now over 70 per cent of preschools are privately run. In 2012, the education department received special funding of up to RMB100 million (A$19 million) to improve the PAP facilities. After consulting with external experts, the department decided to allocate RMB65 million (A$12.4 million) to 12 districts based on the size of the student population and fairness concerns about fiscal capacity. The remaining RMB35 million (A$6.7 million) was allocated based on each district’s proposed performance.

On 29 May 2012, directors (or deputy directors) of the education departments from the city’s 12 districts made a presentation on existing conditions, challenges, vision and strategic goals for their PAP. They also proposed operational plans, detailed budget plans and expected results, in front of a 30-member jury. The jury comprised parents, schoolteachers, officials from the education and finance departments, members of the city’s congress and its People’s Political Consultative Conference and experts, such as university professors. They were also required to address any concerns raised by the jury members.

The jury evaluated the program plans and ranked all 12 proposals. The budget allocations among the 12 districts were based on the jury’s assessment. After the program was finished in March 2013, third-party consultants evaluated the results for each district. Instead of another round

---

6 Private Affordable Preschools are those that provide decent-quality service, but are underequipped. The city requests that such preschools charge a rate that is slightly lower than the market price and promises to subsidise them to upgrade their facilities.
of competitive allocation through on-site presentations, the performance rankings were used for the next year’s budget allocation for the PAP program.

Back to the adoption stage: From late 2012

Due to its efforts, the education department was viewed as the best example of the use of performance measurement in the budgeting system in Guangzhou. The department’s financial officer was invited by the finance department to present some civil servants’ training programs to share the education department’s experience.

Unfortunately, the education department’s passion for PBB did not last long and, since late 2012, it has switched its focus to other budgeting innovations. There are several possible explanations for the change in dynamics.

First, due to a leadership change in the finance department, PBB was no longer prioritised. The director of the department’s PEO also transferred to another agency. In three years, the director of the PEO has changed twice, which is very discouraging.

The education department was not convinced by the new proposals the PEO developed. For an agency that had won the support of top leaders, the education department considered some technical challenges—such as upgrading performance indicators and benchmark analysis—more urgent than the PEO’s proposals at that time. With disappointment, the education department decided to switch its focus to other reforms, such as internal control.

Second, performance evaluation can cause an accountability deadlock, which can discourage line agencies from stepping forward. A key prerequisite for good performance is the ability to spend money. However, in practice, the significantly increased controls on budgeting and the tediously long review and approval process—which is often beyond the control of the line department, especially for basic construction and government purchases—often result in delays or cancellation of programs.

For education services, summer and winter breaks are the best times for construction projects. But, with approval delays, it is impossible for such programs to be fully implemented. Therefore, the agency is stuck with perceptions of poor performance. According to Niu and Song (2015b),
the average implementation period was 504 days for 299 year-long construction and renovation projects—much longer than a fiscal year (365 days). That is, the current performance evaluation saddles the line agencies with responsibility for a burden that is not of their making.

Third, PBB reforms were promoted through administrative orders by the finance department, but it has no legal power over line agencies. And, in any case, budgeting and financial management usually account for only 5 per cent of the total score in the evaluation of departmental performance.

In addition, apart from PBB, the finance department also initiated other reforms, such as internal control and government accounting. The competition for attention among reform programs is fierce. In such circumstances, it is rational for the education department to temper its efforts in adopting performance evaluation. That is not to say that the department disrespects performance measurement. In fact, it still follows the finance department’s guidelines for PBB, such as creating a program inventory and enhancing real-time monitoring of program operation.

It is, however, no longer a pioneer. The impact of performance evaluation results on budget allocations and program management has thereby reduced in recent times.

Conclusion

This study uses the education department in Guangzhou as a case study to explore how a line agency in China can use performance evaluation. Inspired by de Lancer Júnior and Holzer’s (2001) argument about the two stages of performance measurement, this chapter finds that both stages—adoption and implementation—existed in the education department’s case.

From 2006 to 2008, the department mainly followed the guidelines developed by the finance department, as did other agencies, without further innovation to improve result-oriented budgeting reform. However, from 2009, due to a leadership change and external pressure, the education department became ambitious and actually implemented, not just adopted, the PBB instrument. It did this by extending the scope of evaluation, inviting experts to help build organisational capacity,
launching pre-program evaluations and welcoming citizen involvement. All these endeavours made the education department the pioneer of line agencies in Guangzhou in promoting performance measurement.

Surprisingly, however, after three years’ hard work, the education department stopped taking the lead and went back to the adoption level, involved only in those activities required by the finance department. This case analysis finds that disagreement between the finance department and the education department about reform strategies was the major reason for the latter making performance measurement less of a priority. The approach to promoting PBB reform and the administrative processes involved (with long time frames) caused conflicts between the existing program management system and the accountability mechanism under PBB, discouraging the education department from continuing its pioneering innovation. It is evident that a shared understanding between the finance department and line agencies is essential for successful implementation of PBB. This includes the finance department acknowledging the expertise and administrative experience of the line agencies. Moreover, it seems to take time for China to incorporate the PBB instruments into the existing accountability system, which is the fundamental challenge and explains why the performance results have a very limited impact on decision-making.

De Lancer Julnes and Holzer (2001) discovered that rational factors have a preponderance of influence on the adoption stage and political factors have a preponderance of impacts on the implementation stage. This does not seem to be an inevitable conclusion in China. The swing between adoption and implementation of PBB in the education department shows that, although the rationale for performance measurement was appreciated, the contextual factors, including both political and organisational variants, are vital for making the reform sustainable.

Ho and Im (2015) argue that applicability and appropriateness should be carefully examined when a Western-oriented reform is borrowed for use in a developing country. Although it is debatable whether China is a developing country, the key point is that the Western-style reform established on a technical rationale will have to adapt into non-Western contexts to achieve the dynamics of the reform. This case study suggests that the next design for PBB reform must not only take the technical upgrading into account, but also develop strategies to deal with institutional barriers.
References


