Corporate Level Performance Measurement in Tax Administrations in Emerging Economies: Towards Convergence?

Elizabeth Kariuki

July 2012

www.april-ssa.com
Corporate Level Performance Measurement in Tax Administrations in Emerging Economies: Towards Convergence?

- Elizabeth Kariuki

Keywords: case study, fuzzy set logic, key performance indicators, performance measurement, qualitative comparative analysis, revenue authorities, semi-autonomous revenue authority, strategic plans, tax administration

Background and context

In the two decades from 1985, many emerging economies have restructured their tax administrations by merging former income tax and customs and excise departments within their Ministries of Finance, and hiving them off into agencies that are commonly referred to as semi-autonomous revenue authorities (ARAs). It is suggested that the creation of ARAs was influenced by the new public management cultural paradigm, and cultivated by institutions such as the World Bank, International Monetary Fund (IMF), Organization for Economic Cooperation and Development (OECD), and the UK's Department for International Development ((Fjeldstad and Moore, 2008: 2); (Therkildsen, 2004: 75)).

There are also claims that ARAs are comparable to central banks which constitute “national financial institutions [that] are relatively transferable” (Fjeldstad and Moore, 2008: 9). In this regard, for example, it is noteworthy that many ARAs across the globe are established by legislation. Furthermore, under their enabling legislations, ARAs have varying degrees of autonomy to discharge their responsibilities on a self-financing basis as well as, hire and remunerate staff outside the Public Service rules and regulations. Moreover, ARAs are less prone to political intrusion (Devas et al., 2001).

One of the IMF’s prescriptions to do with reforming an ARA is that it should “diagnose the…principal problems, design a strategy for actions to be taken, and agree on…scope and timing” (Silvani and Baer, 1997: 12). This remedy is reinforced by the OECD (2001: 3) which suggests it is good practice for ARAs to be “strategically focussed and responsive to changes in their environment and that of their taxpayers”. Therefore, corporate level strategic plans, which contain a set of goals (outcomes) and outputs, as well as associated performance measures and targets, is a central feature of ARAs’ management processes.

The outcomes and outputs sought are typically influenced by an ARA’s “mission; core functions and capabilities; [the types of ] stakeholders/
[taxpayers]; and stakeholders’/taxpayers’ expectations” (Cole and Parston, 2006: 67). Their associated key performance indicators (KPIs) should guide ARAs in gauging whether outcomes and outputs have been achieved. However, in practice, it is not obvious why ARAs select certain KPIs and not others. Against this background and context, this paper seeks to disseminate the results of exploratory research on: Which factors influence the selection and use of performance measures? Furthermore, on the basis of the research findings, the paper proposes a menu of KPIs which can be used by ARAs in emerging economies to assess the effectiveness of their operations, and for cross-country benchmarking.

**Overview of methodology**

The methodology, findings and conclusions are drawn from this author’s doctoral research. The methodology used comprised eight steps. The first step entailed a select review of the literature on strategic planning, performance measurement and information use in a public sector setting, covering the work of academics such as: (Hood, 1991); (Smith, 1995); (Propper and Wilson, 2003); (Pollitt and Bouckaert, 2004); (Bevan and Hood, 2005); (Chun and Rainey, 2005); (Pidd, 2005); (Behn, 2006); (Neely, 2007); (Talbot, 2010); (Boyne and Walker, 2010); and (Boyne, 2010).

The second step involved a review of corporate plans and annual reports of 10 ARAs from around the world to discern the broad outcomes sought, and any common KPIs. In this regard, the review identified 11 broad outcomes which the 10 ARAs use on a selective basis (see Table 1). Third, the author pinpointed universally available output and outcome KPIs developed by academics, in particular the work of Vázquez-Caro and Bird (2011), and institutions such as the World Bank and United States Agency for International Development (USAID), that would be a suitable match for the eleven common broad outcomes identified.
Corporate Level Performance Measurement in Tax Administrations in Emerging Economies: Towards Convergence?

Table 1: The 11 broad outcomes by key result area

<table>
<thead>
<tr>
<th>Key result area</th>
<th>Broad outcome (O)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Revenue collections</td>
<td>• O1 – Increased revenue collections for development</td>
</tr>
<tr>
<td>2. Modernisation</td>
<td>• O2 – Operations modernised through automation</td>
</tr>
<tr>
<td>3. Taxpayer services</td>
<td>• O3 – Improved taxpayer services</td>
</tr>
<tr>
<td>4. Compliance</td>
<td>• O4 – Taxpayers find it easy to pay taxes and hence comply</td>
</tr>
<tr>
<td>5. Enforcement</td>
<td>• O5 – Deliberate non-compliance effectively addressed through enforcement</td>
</tr>
<tr>
<td>6. ARA costs</td>
<td>• O6 – The ARA operates efficiently &amp; effectively</td>
</tr>
<tr>
<td>7. Protection of borders and trade facilitation</td>
<td>• O7 – Size of the illicit economy minimised</td>
</tr>
<tr>
<td>8. ARA internal capacity</td>
<td>• O8 – Legitimate trade enhanced</td>
</tr>
<tr>
<td>9. ARA work environment</td>
<td>• O9 – Staff capacity /or administrative systems strengthened</td>
</tr>
<tr>
<td>10. Good governance</td>
<td>• O10 – An enabling work environment in place</td>
</tr>
<tr>
<td></td>
<td>• O11 – An ARA with a strong corporate image and which is well governed</td>
</tr>
</tbody>
</table>

Source: Kariuki (2012)

The fourth step built on the preceding steps to develop a questionnaire and interview guide for internal stakeholders (ARA executives), and a topic guide for external stakeholders (academics, tax practitioners and civil society organisations). Fifth, the author administered the data collection instruments among 51 ARA executives and 17 external stakeholders in three countries – Kenya, South Africa and Tanzania. The results of the survey of ARA executives and external stakeholders were used two identify two sets of 12 priority KPIs. Sixth, the feedback from qualitative interviews with ARA executives as well as institutional documents provided a basis for developing country case studies for Kenya Revenue Authority (KRA), South African Revenue Service (SARS), and Tanzania Revenue Authority (TRA).

The seventh step involved the identification of the factors that influence the adoption of performance measurement in ARAs. This involved a further analysis of theory and comparing it with the evidence contained in each ARA’s case study. This author drew on fuzzy-set techniques for further in-depth analysis and to make generalisations ((Ragin, 2000); (Ragin, 2008)). In this regard, based on the available evidence, an assessment was made on the degree of set-membership of the three case studies in eight causal conditions. The range of membership ratings applied were as follows: (1) fully in the set= 1; (2) more in than out of the set = 0.75; (3) neither in nor out of the set= 0.5; (4) more out than in of the set=0.25; and (5) fully out of the set= 0 (Ragin, 2000: 156). The final step entailed using qualitative comparative techniques (QCA) (such as fuzzy-set algebra, configurations and truth tables), to identify prominent causal conditions and combinations of conditions that lead to the adoption of performance measurement in ARAs. A freely available software package, fuzzy-set QCA (fsQCA), was applied to aid the analysis (Ragin et al., 2008).
Key findings and some recommendations

The research found that there are probably three key findings that provide plausible answers to the areas of enquiry posed at the beginning of this paper. The first finding covers the causal conditions (independent variables) influencing the degree of adoption of performance measurement in ARAs (the dependent variable (Y)). The second finding further analyses the causal conditions as a means of pinpointing their interrelationships, and which ones appear most dominant. The third finding centres on the results of the review of the literature and feedback from both ARA external and internal stakeholders as a means of generating an indicative menu of KPIs for each broad outcome.

Conditions influencing degree of adoption of performance measurement

Degree of adoption of performance measurement

The survey of ARA executives identified 12 priority KPIs out of a total of 39. The priority KPIs cross-referenced to the broad outcomes in Table 1 are as follows: (1) actual revenue compared to forecast revenue (O1); (2) voluntary compliance rate (O4); (3) percentage of taxpayers satisfied with services/information and tools provided by the ARA (O3); (4) percentage uptake in electronic filing (O3); (5) average processing turnaround time for tax returns and refunds (O3); (6) reduction in taxpayer/trader compliance burden (O4); (7) amount spent for every dollar collected (O6); (8) percentage increase in accuracy of processing (O6); (9) time taken to prepare, file and pay various taxes (O4); (10) global ease of paying taxes ranking (O3); (11) percentage uptake in electronic filing (O3); and (12) number of cases successfully prosecuted (O5).

The 12 priority KPIs were used to determine set membership in the dependent variable (Y) – the degree of adoption of performance measurement in ARAs. The degree of membership was assessed by ascertaining from the case studies how many of the 12 priority KPIs have been applied in each ARA over time. The degree of membership was set as follows: (1) nine or more KPIs adopted = 1 - full membership of the set; (2) more than six but less than nine KPIs adopted = 0.75 - more in than out of the set; (3) three up to six KPIs adopted= 0.5 - in between; (4) one or more but less than three KPIs adopted = 0.25 - less out than in of the set; and (5) no KPIs adopted = 0 - not a member of the set (see Table 2).
Table 2: Membership of the set “adoption of performance measurement in ARAs”

<table>
<thead>
<tr>
<th>ARA</th>
<th>Set Y – The degree of adoption of performance measurement in ARAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS</td>
<td>1</td>
</tr>
<tr>
<td>TRA</td>
<td>0.75</td>
</tr>
<tr>
<td>KRA</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Key: 1= nine or more KPIs adopted; 0.75 = more than six but less than nine KPIs adopted; 0.5 = six KPIs adopted; 0.25 = three but less than six KPIs adopted; 0 = no KPIs adopted

Source: Kariuki (2012)

Membership in the eight sets and their influence on adoption

On the basis of the theory on performance measurement, the eight causal conditions identified were: (1) pragmatism is a dominant political economy driver in setting performance measures (A); (2) good practices influence the choice of outcome KPIs (B); (3) target setting is a chief feature of performance measurement (C); (4) policy and legislative requirements make a significant difference to the performance measurement regime (D); (5) the structure of the revenue base influences the choice of KPIs (E); (6) a performance culture has a bearing on the selection of KPIs (F); (7) priority information demanded by external stakeholders impacts the choice of performance measures (G); and (8) data collection cost and quality are taken into account in selecting KPIs (H).

The degree of membership of the three case studies (KRA, SARS and TRA) in the eight conditions, suggests that all factors play a part in the adoption of performance measurement in ARAs. Table 3 indicates their varying memberships. It also presents minimum and maximum overall membership scores using the fuzzy-algebra equations ‘and’ (∩) and ‘or’ (U) to combine all conditions.

Table 3: The eight conditions shaping performance measurement in ARAs

<table>
<thead>
<tr>
<th>ARA</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Cralln</th>
<th>Crallu</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARS</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>TRA</td>
<td>1</td>
<td>0.25</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.75</td>
<td>0.75</td>
<td>0.75</td>
<td>0.25</td>
<td>1</td>
<td>0.75</td>
</tr>
<tr>
<td>KRA</td>
<td>0.75</td>
<td>0.25</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
<td>0.25</td>
<td>0.5</td>
<td>0.25</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Key: A= Pragmatism is a dominant political economy driver in setting performance measures; B= Good practices influence the choice of outcome KPIs; C= Target setting is a chief feature of performance measurement; D= Policy and legislative requirements make a significant difference to the performance measurement regime; E= The structure of the revenue base influences the choice of KPIs; F= A performance culture has a bearing on the selection of KPIs; G= Priority information demanded by external stakeholders impacts the choice of performance measures; H= Data collection cost and quality are taken into account in selecting KPIs; Cralln = Minimum score (A ∩ B ∩ C ∩ D ∩ E ∩ F ∩ G ∩ H); Crallu = Maximum score (A U B U C U D U E U F U G U H)

Source: Kariuki (2012)
Causal conditions with the greatest influence on the adoption of performance measurement

Table 3 suggests that causal condition C (target setting) is dominant in terms of theoretic consistency and coverage (see Table 4). In fact analysis using fsQCA indicates that all other conditions are a sub-set of C — meaning that they must be present in order for C to occur. In reality, this is a plausible explanation as other factors determine what targets are fixed. Moreover, condition E (structure of the revenue base) is the next dominant factor – again, given that ARAs are mandated to administer various types of taxes, this finding can be expected. The least dominant causal condition is B (use of good practices in setting outcome KPIs). Considering that the application of outcome KPIs in ARAs is still in the early stages, this result is consistent with current practice.

Table 4: Set membership in outcome C (target setting)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Raw coverage</td>
<td>0.75</td>
<td>0.5</td>
<td>0.75</td>
<td>0.83</td>
<td>0.75</td>
<td>0.67</td>
<td>0.75</td>
</tr>
<tr>
<td>Combined</td>
<td>0.91</td>
<td>0.79</td>
<td>0.91</td>
<td>0.94</td>
<td>0.91</td>
<td>0.87</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Notes: Theoretic consistency is defined as “the degree to which solution terms and the solution as a whole are subsets of the outcome” – in this case C. Theoretic coverage is defined as “how much of the outcome [C] is covered (or explained) by each solution term and by the solution as a whole” (Ragin et al., 2008: 85).

Source: Kariuki (2012)

Further analysis of combinations of the conditions (or configurations), and combinations of configurations in fsQCA suggest that ‘configuration Comb3=Minimum score (E ∩ F ∩ H)’ has the highest coverage. In fact, when combined with other configurations (e.g. C ∩ E or C ∩ D or A ∩ C ∩ E), the configuration of conditions E (structure of the revenue base), F (performance culture) and H (data cost and quality) account for 25% of the instances of the occurrence of outcome ‘Y’. In contrast, the configuration of the condition ‘Tratop3’ comprising A (being pragmatic), C (target setting) and E (structure of the revenue base) only embraces a mere 10% of the incidence of outcome ‘Y’. The remaining 65% represents an overlap.

The production of a truth table using fsQCA software should generate possible combinations of conditions for the occurrence of outcome Y – in this case given that k=8 (causal conditions), the truth table should generate 256 combinations. When this author attempted to generate a truth table, because the number of cases was so small, many of the results were ‘remainders’ (blanks), with no empirically supported evidence. Nevertheless, based on the three cases, the truth table suggests that the most complex solution to ‘Y’ with sufficient coverage and consistency is ‘C ∩ E ∩ F ∩ H’, where as ‘E ∩ F ∩ H’ is a more parsimonious solution.
The complex solution above, which includes condition ‘C’, resonates with other empirical research which reports that target setting is crucial. Specifically, according to Boyne (2010: 222), when a KPI is assigned a target, it has a “positive effect” on performance. In addition, given the other three conditions (E, F and H), it seems essential to bear the following points in mind when selecting KPIs: (1) consider the measures that best capture strategic initiatives by type of tax and/or taxpayer (to cater for condition E); (2) KPIs are likely to be refined as the institution learns and as a result improves – these considerations underpin a performance culture (Marr, 2009); and (3) where appropriate, it makes sense to take advantage of readily available KPI data such as universal indicators as a means of containing costs.

**Towards convergence – a potential set of standard measures**

ARA executives as well as external stakeholders selected priority KPI’s for the broad outcomes listed in Table 1. In this regard, it is striking that the research found that as long as an ARA’s mandate is to collect revenues to meet public expenditure, ‘actual against forecast revenue’ is likely to not only be the top KPI, but will also influence the choice of at least one corporate goal. Furthermore, on the whole, external stakeholders within the three jurisdictions consider the majority of commonly used performance measures to be relevant. Still, there is also demand across the board for ARAs: to report on measures particularly to do with equity, internal capacity, corruption and tax expenditure; and to benchmark performance data with other tax authorities.

The feedback from internal and external stakeholders was enriched by the ongoing discourse by academics on how best to benchmark ARAs’ performance. For instance, concerns raised by external stakeholders around the importance of ARAs reporting on equity considerations was catered for by drawing on the work of Vázquez-Caro and Bird (2011), who concur that tax administrations should keep tabs on it. To this end, as none of the 10 ARAs covered in the literature review explicitly address, equity, this author recommends its inclusion as a twelfth broad outcome – ‘the tax system is equitable and fair’.

**Table 5** sets out a potential set of standard KPIs based on the three sources listed above. ARA executives are of the view that there is utility in adopting the following two universal KPIs developed and administered by the World Bank: “time taken to prepare, file and pay various taxes”; and “global ease of paying taxes ranking”. One universal KPI that respondents identified as unsuitable is the ‘tax staff per population’ (developed by USAID). Therefore, the three replacement KPIs included for ‘O9 – staff capacity and/or administrative systems strengthened’ are drawn from interviews with ARA executives, SARS’s current strategic plan and the work of Vázquez-Caro and Bird (2011).
# Table 5: A possible set of KPIs

<table>
<thead>
<tr>
<th>Broad outcome</th>
<th>KPI</th>
</tr>
</thead>
</table>
| O1 – Increased revenue collections for development | • Actual revenue compared to forecast revenue (in aggregate and by tax type)^<sup>+</sup>  
• Tax gap<sup>*</sup>  
• Revenue growth projections<sup>*</sup>  
• Tax expenditure/ exemptions<sup>*</sup> |
| O2 – Operations modernised through automation | • % uptake in eFiling<sup>*</sup>  
• Time taken to prepare, file and pay various taxes<sup>*</sup>  
• No (% of transactions with internet access>  
• No (% of totally automated interactive transactions> |
| O3 – Improved taxpayer services | • Reduction in taxpayer compliance burden<sup>*</sup>  
• Average processing turnaround time for tax returns and refunds<sup>*</sup>  
• % taxpayer satisfaction with services/information & tools provided by ARA<sup>*</sup>  
• Time taken to prepare, file and pay various taxes<sup>*</sup>  
• Global ease of paying taxes ranking<sup>*</sup> |
| O4 – Taxpayers find it easy to pay taxes and hence comply | • No of (or %) cases recommended for prosecution/ successfully prosecuted<sup>+</sup>  
• % of exceptions uncovered from the risk engine<sup>+</sup>  
• No (or %) of non-compliant cases uncovered as a result of audits<sup>+</sup>  
• Average number of days taken to identify stop-filers<sup>+</sup> |
| O5 – Deliberate non-compliance effectively addressed through enforcement | • No of (or %) of cases recommended for prosecution/ successfully prosecuted<sup>+</sup>  
• % of exceptions uncovered from the risk engine<sup>+</sup>  
• No (or %) of non-compliant cases uncovered as a result of audits<sup>+</sup>  
• Average number of days taken to identify stop-filers<sup>+</sup> |
| O6 – The ARA operates efficiently & effectively | • Amount spent for every dollar of tax collected or administrative costs as a % of revenue collected<sup>+</sup>  
• % increase in accuracy of processing<sup>*</sup>  
• % reduction in escalated service queries<sup>+</sup> |
| O7 – Size of the illicit economy minimised | • % decrease in the size of the illicit economy<sup>*</sup>  
• Customs compliance index<sup>*</sup> |
| O8 – Legitimate trade enhanced | • Reduction in trader compliance burden<sup>*</sup>  
• % trade volume coverage by preferential traders<sup>*</sup>  
• Percentage increase in the number of AEOs<sup>*</sup> |
| O9 – Staff capacity and/or administrative systems strengthened | • Tax staff as a proportion of registered (or active) taxpayers<sup>+</sup>  
• Productivity per employee<sup>+</sup>  
• Proportion of staff in key managerial, technical and professional positions<sup>+</sup>  
• Training impact on tax administration performance<sup>+</sup> |
| O10 – An enabling work environment in place | • Employee satisfaction rate or index<sup>*</sup>  
• Training impact on tax administration performance<sup>+</sup> |
| O11 – An ARA with a strong corporate image and which is well governed | • Governance/ corruption indicators (e.g. East Africa Bribery Index* and Transparency International’s corruption perception index*)  
• Unqualified audit report<sup>*</sup> |
| O12 – The tax system is equitable and fair | • Level of horizontal equity<sup>+</sup>  
• Level of vertical equity<sup>+</sup>  
• Number (% of administrative general rulings conceived collectively with civil society<sup>+</sup>  
| Source: Kariuki (2012)
Concluding remarks

This paper, which draws on this author's doctorate research, contributes to the discourse on performance measurement in ARAs in emerging countries in two fundamental ways: it provides insights around the conditions affecting the adoption of performance measurement; and identifies a potential set of standard measures. Both contributions have the potential to lead to the convergence of KPI use in ARAs beyond the conventional reliance on financial metrics to tell the whole story about performance, especially those evolving around revenue collections. In this latter regard, the paper demonstrates that on the basis of outcomes that meet their contexts, ARAs can purposively select measures (particularly those that are commonly used and universally available).

However, the small number of cases used in the research means the diverse conditions that may be present in a wider group of up to 185 ARAs globally have not been tested. This is one area worthy further exploration – in other words expanding the number of country cases would facilitate more robust theory-building around the eight causal conditions. There is also scope for providing clearer definitions and methods for measuring some of the KPIs such as: 'training impact on tax administration performance'; and horizontal equity, which needs to be disaggregated by type of tax.
Bibliography


