參、專題演講 II
Keynote Speech II

摘要：

公部們使用績效評估是一個非常具有爭議性的議題，績效評估跟目標有關，其評價兩極，而績效評估面臨到兩個問題，一個是沒有辦法去評估重要的事情，另外一個就是必須在工作的時後設定目標。在公共行政的制度之下，公務人員應該要很務實，不應太過理想化，公部門碰到問題時，則要有批判思考的精神，不能憤世嫉俗，而是找出解決方案。大家對績效評估的期待，可以讓我們用績效評估的方式對人員來做評量，再來就是透過績效評估來做爲如升遷、加薪的工具。然而績效評估面臨到一些問題，分別是 (1) 績效評估的方式太過簡單，只看投入的金錢和時間，以及最後的產出，而未在績效評估的時候去評量結果跟過程。因此，我們不知道長期下來某些政策的結果或是表現的影響力在哪裡？ (2) 績效評估只有去看個別的表現，並沒有把大環境納入考量 (3) 我們的資源受限制，好像沒有辦法再達成目標 (4) 沒有把歷史脈絡納入考量，不能只看現在的績效爲何沒達到目標。由於有以上問題，因此績效評估是否適用到公部門爲探討之重點，我們訪談的公務員，認爲要透過績效評估的方式讓外界看到，公部門其實做得非常好，重點不是要解決問題，問題在於沒有一個工具、一個數據給大家看。讓社會大眾從這個角度看到公部門，必須要有很好、可靠的績效指標，但是需要考量到所有會影響到相關績效的因素，重點就是環境、情境，請務必要把環境、情境、相關條件放入，如此才能夠得到最正確的答案。

貳、專題演講 II

一、時　間：101 年 10 月 15 日 (星期一) 下午 1 時 30 分

二、主　題：How to use performance measures and to convince public officials that they are indeed useful.

三、主持人：詹中原考試委員

四、演講人：Michiel S. de Vries 國際行政學校暨機構聯合會主席

五、內　容：

非常謝謝本次會議的主辦單位，我很榮幸能夠參與這次的會議並且擔任專題演講人。首先，我不知道大家對我要談的內容同意與否，無論如何我覺得有不同觀點是好的，因為我喜歡討論。其次，如果大家對我所說的有不同見解，我希望能夠有些討論、辯論，這樣會更好。

我們要談的是公部門使用績效評估，這是一個非常具有爭議性的議題，因為我們講到績效評估的時候就跟目標有關係，要去評估是否能達到目標。私部門使用績效評估越來越普遍，也愈來愈多人支持績效評估，其認為績效評估可以讓一般人員去達到目標，但也有很多人持保留態度，不知道績效評估是否眞的能夠達到所謂的效果？是否眞的能夠反映出大家在工作上的表現？這也表示績效評估是否眞的有所謂的效度？有許多學者認為公共行政非常複雜，無法用簡單的績效評估去判定是否達到目標，因為它無法完整的反映出公務體制的效度，因此，部分學者主張要用不同的方式做評估會比較適合。

有些醫療相關的研究報告指出，在做績效評估的時候反而會影響到醫療人員所提供的醫療品質，亦即績效評估並無法讓這些醫療人員去達成良好的服務，反而會影響病患所獲得的服務品質。常言道：「有時候我們沒有辦法去評估重要的事情」，這個就是績效評估所面臨的問題；另外一個績效評估所面臨的問題就是我們必須在工作的時候設定目標，例如：我們必須在 1 個月內回應民眾的來信，如果在期限內我們只回了 90% 的信，似乎就沒有達到目標，這時政府就會說：「的確沒有達到這個目標，可能是因為我們的業務量太大了，又或者是說我們面臨太多的問題，所以表現不如預期，也不是我的問題。」。

在公共行政的制度之下，諸位都是公務人員，所以我們應該是務實的，而非理想化的，理想化這樣的事情是要讓政治人物來做的，我們必須要讓公共服務是很務實、很實際的，而我們也希望能有好的績效表現。面對問題，我們應該要有批判思考的精神而不是憤世嫉俗，試著找出解決方案，這就是我們希望達成的目標。接下來會談到績效評估所面臨的問題，其亦為世界各地公部門所面臨的問題。目前很多人都不支持績效評估，所以要怎麼樣讓績效評估能夠做得更好，要怎麼樣去提升它的運用，即為接下來要談的重點，讓我們來談談大家對績效評估有哪些期待，如中途有不同意見，都可提出來做討論。

績效評估是讓我們可以用評量的方式來對我們的公務人員作考核，比如說有時候你覺得一個人績效沒有達到標準，那麼就要被解雇，這個是大家的觀念或思維；再者就是可利用績效評估來做部

門控制，比如說某些部門做得不好、表現不盡理想，就應裁撤該部門；或是透過績效評估來激勵我們的公務人員，比如說有些人的表現超乎預期，本來預定目標爲 80 分，但他做到 100 分，此時，就可以利用加薪等方式來激勵我們的公務人員；還有績效評估也可用來做爲升遷的依據，表現不錯的人就可以更上一層樓，而且可以從中互相學習，比如說，某個部門表現良好，就可成爲其他部門的模範並學習。所以績效評估可以促使大家精進自己的工作能力，如能確實完成目標者，則能讓公部門提升績效，並讓我們知道每一個部門的工作表現及公務人員做事的方法正確與否、投入的經費和時間是否妥適、如何激勵員工、如何說服其他人了解到我們的組織表現良好，以及如何發現做得好及表現好的業務單位、還有我們哪些做法不錯、哪些做法是行不通的，如果某些措施行不通，要如何改進並提出的解決之道。

雖然說績效評估能夠讓公部門更好，但仍有漫漫長路需要去改進，因爲很多表現不好組織單位要去改變某些做事的方法是很曠日費時的，即便我們知道說要如何改進，但改進的方法我們不曉得，也不知道要如何進行，也沒人知道改善問題的方法在哪裡，而改革就像是一個賭注，對於實際改革的人員來說是冒很大風險的，這個時候公務人員的士氣就會下降，因爲我們不知道應如何改善現有的做法，也不知道解決的方法在哪裡，有的人聲稱知道改變的方式，但很多時候是高估了自己，或許大家有些不同的意見，但這就是我的看法，所以我們必須要去找到如何解決和如何說明現有狀況的方法。

而我自己的做法是什麼呢？我覺得必須要瞭解政府的績效表現是經由有哪些因素所決定的，然後再解釋爲什麼某些單位表現的不錯，某些單位表現的不盡理想，並且要協助政府運用現有的資訊，我們必須要有好的做法說服公家機關的人員，績效評估對他們來說是有利而無害的，因爲有很多人會認爲之所以要做績效評估是因爲公家機關的績效不好，所以才要做，但這只是一個基本的假設而已。

很多人認爲現在政府部門的運作有問題，所以讓很多公務人員覺得自己表現不好，進而覺得很難爲情，而我覺得這是很多國家的思維，我覺得應該以自己能在公部門服務爲榮。雖然公家機關的業務非常的複雜，但大部份人都表現得很好，雖然說民眾的標準可能很高，抑或期待也很高，但就公部門的表現或績效來說是好的。我覺得反而是要用績效評估來反轉大家的觀念，扭轉大家對公部門一些排斥的情緒，要讓大家看到我們的表現其實很不錯，我們單位的運作是很良好的，不要把績效評估當成一個敵人而是要當成一個朋友。

但是現在面臨的問題，各領域的學者專家皆有不同的論點及觀點，有的人會覺得重點不是你做什麼，而是你的表現或你做事的方式，或者是說政策是否眞的能夠促成改變。所以績效評估所面臨的問題到底是什麼呢？就是績效評估的方式太過簡單，我們常常說你的結果到底是什麼，或者是它占的比重到底有多少等等，又或者是我們在審視績效評估的時候，只會去看所投入的金錢與時間，還有最後的產出，比如說我們回應多少民眾的信件，或者是我們又提供了多少的補貼，但是我們沒有在績效評估的時候去評量所謂的結果跟過程，這就是一個問題。我們不知道長期下來某些政策的結果或表現的影響力到底在哪裡？除此之外，績效評估並沒有將大環境納入考量，只單單去看個別的表現，比如說整個組織所在的大環境所受到的影響就像是現在歐洲所面臨的金融危機一樣，而我

認爲現在的臺灣也應該是如此吧！

另外，我們所面臨的問題就是有限的資源使我們無法達成目標，但這並非是我們的問題，可是其他人會認爲這就是我們的問題，因爲我們事前沒有將大環境的變數納入考量。此外，在做績效評估的時候，也沒有將歷史因素納入考量，實際上，我們必須要考量歷史的脈絡，即過去到現在的發展是如何演變，而不能只看現在的績效是否達成，我們要從歷史的脈絡來探討，再來看我們衡量的指標和政策之間的關係。還有，我們的想法應該是要去衡量結果，但到底是什麼東西的結果呢？事實上，我們很難去找出在績效評估方面最好的作法，因爲要面臨這麼多的困難，這也難怪公務人員只要講到績效評估就會有點遲疑。

首先，我們要問的是績效評估到底適不適用於公部門，有些人會認爲不適用，因爲他們認爲績效評估沒有辦法發揮功能，反而容易造成意料之外的負面效果，這更無法使政府充分運作，這是美國某個知名的學者所提出來的論點，而我的研究是有關荷蘭官員，荷蘭政府官員跟臺灣的政府官員很像，也就是說，我問的是這些公務人員本身對於在公部門運用績效評估有何看法，還有如果他們意見不同的話，我們要如何解釋爲何有如此的情形發生，所以我這邊的資料庫是來自我們的內政部，它的資料非常豐富，從地方政府、省政府甚至到中央政府、警察局、大學及醫療人員都列入在這個資料庫裡頭，因此資料來源是非常龐大的。

在這份問卷調查當中，其中有個問題就是你的工作是不是可以衡量？是不是喜歡有人來衡量你的工作績效？在 27,000 多名的公務人員當中，其中 50% 的人認爲是可衡量的，另有 50% 的人認爲是無法衡量的，這樣的情境對我們做研究的人來說，就很有趣了，這樣子我們就可以想辦法來解釋爲什麼大家有如此兩極的看法，有些人認爲績效評估可以衡量他們的工作內容，有些人則認爲不行，在這邊提出幾種不同的模型。

第一種模型，有些人會認爲在公部門做事，工作內容太過複雜，所以無法衡量；另外一個模型是所謂的理性模式，假設我們看到一個問題，我們就應該要來衡量它的相關結果，這樣我們才能更知道現況然後加以改進；第三個模型就是所謂的公共選擇模式，也就是我們要來選擇做衡量的標的，假設我現在的職能非常的複雜，那麼我的工作就一定要有人來衡量，因爲我要讓大家看到我的工作有多麼複雜，這就是所謂的第三個模型，在這三種模型之下就會衍伸出三種不同的期望。

還有我們也會問他們有關於同事、同儕的績效問題，比如我們會問他說，你的同事是否會謊報工作績效？或我的同事績效是否不好？如果是這樣的情況，我們是不是需要透過績效評估來找出這樣的問題？有些人會說我們不需要用這樣的方式來保護我們機構，但也有些人的觀點是，我們的績效非常好，我們的部門也非常的好，我們一定要透過績效評估讓世界看到我們工作的成果，經由以上的推論，我們在這邊就可以明確知道爲什麼有 50% 的人認爲他們的工作內容是可衡量的。

但同時也有半數的人不同意，那我們要如何解釋這當中的差異呢？如果公務人員符合這些條件的話，他們就會覺得這樣的績效評估是好的，但要有哪些條件呢？首先，他們的工作內容比較複雜，比較不屬於單純的工作內容；其次，如果他們是在一個複雜的環境下工作且對於同儕的績效是有正面評價的話；最後就是假設他們是管理職，而非政策的相關職位。符合以上這幾個條件，他們就會

覺得績效評估是比較好的，而他們也會想要透過績效評估的方式來保護自己的位置。

要告訴對他們有一點點敵意的外界，透過績效評估的方式證明給他們看，我們的單位做得很好，我們的單位很有價值，不知道臺灣是不是也是這樣的情況，大家是不是都會批評官僚，大家都說你是公務人員，你們一定做得很差、做得很不好，好像大家都不會以在公部門任職爲榮，好像大家都會覺得公務人員比較不聰明、教育程度比較不好等等，也就是說不好的人才會去公部門任職，這是荷蘭某些部份人的看法。

有些我們訪談的公務員說要透過績效評估的方式來讓外界看到公部門其實做得很好，這是必要的，透過這樣的方式，我們才能夠讓他們知道，我們做得好，但這些重點並不是在解決問題，而眞正的問題是在於我們並沒有一個明確的工具或一個明顯的數據給大家看、讓大家了解，公部門是如何運作，更不能說我就是個官僚，說我做得不好，事實上，我做的可好呢，你看這是我們的報告，這就是我們的成績單，這個報告可以看得出來我們的成果，我們的績效眞正了不起，我們對社會很有貢獻。

讓社會大眾可以用這樣的角度來看待公部門，但要確實做到這一步的話，我們就必須要有很好的績效指標，而且必須是信賴可靠的，那我們要怎樣改進這些績效指標呢？因爲現在只有衡量輸入跟輸出，並沒有衡量過程跟結果，也沒有看目標是否已達成，更沒有衡量公務體系的廉潔或永續等等的問題，這些我們都沒碰觸到。目前我們有哪些解決方案呢？我們有哪些量化的指標？你可以問的是你做你的工作是否開心？或是說你是不是覺得你的工作很好？這是我們目前現有的問法，但我始終覺得問得不夠好，有些人會問說，若我們有這個輸出，然後再透過一般的推論，就可以說，因爲我們有這樣的產出，所以我們達成什麼樣的結果，然後我們可以重新定義結果，結果就是長期輸出之後達成的。

我個人觀點認爲這些論點都不夠好，而且我們現在有很多所謂的評鑑委員會，評鑑這個、評鑑那個，然後大家都在談論所謂的最佳做法、最佳實務。實際上，我想要跟大家說的就是，我們必須要考量現實環境，比如說你在評估績效的時候，你也必須要考量到所有會影響到有關績效的因素，那是因爲需要研究報告來佐證，所以你需要大學這樣的單位來幫忙。

比如說我們想要降低嬰兒死亡率，或者是我們想要提高識字率，那你就要問說，識字率會受到哪些因素的影響呢？有哪些因素會取決於識字率的高低？當然是視一個國家的發展程度，這跟公部門的績效沒有任何關係，它只是其中的一種因素而已，再來就是都市化程度，一個國家都市化程度，比如說在城市裡的學校就比在村落裡多，所以識字率就會高，再來還有就學率，有多少孩童確實有上學，那你就可以考量這些因素，這些因素可以解釋識字率，當然其中也有一些不能解釋的部份。

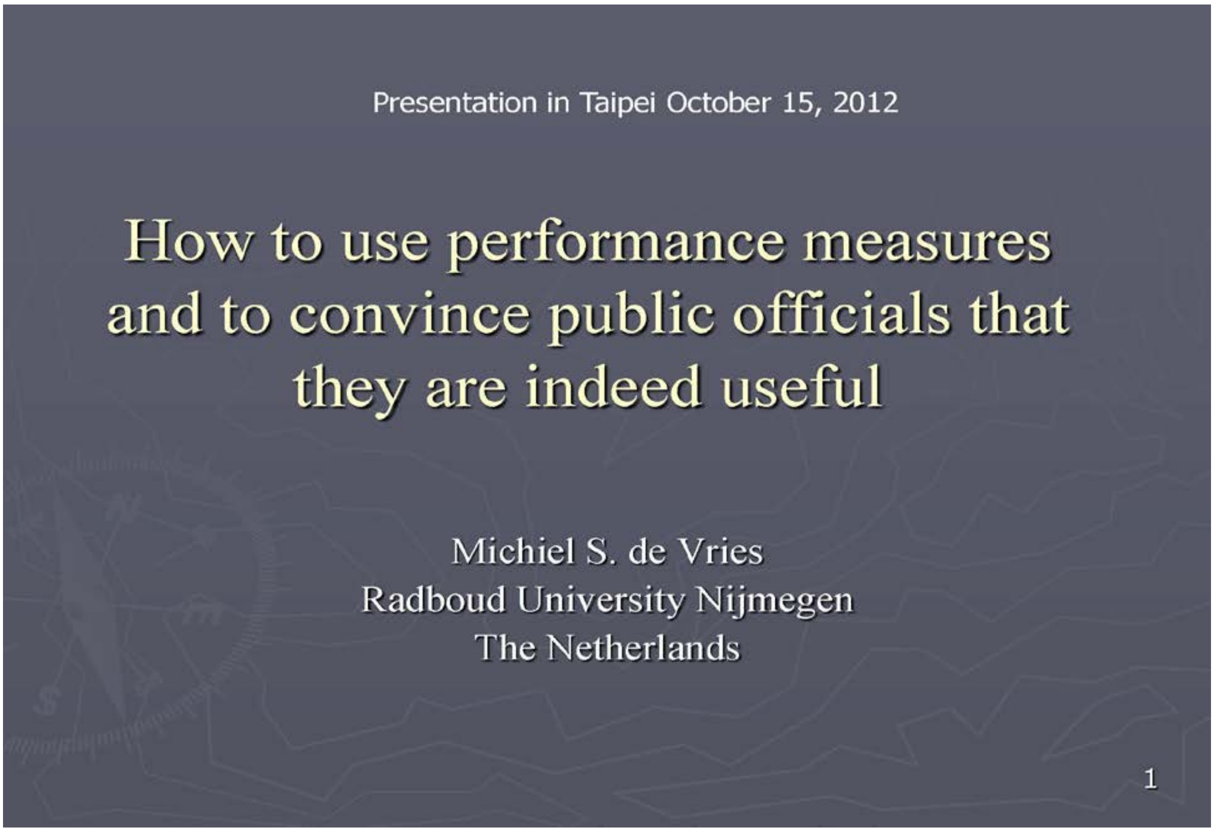
透過迴歸分析把這些獨立變數跟應變變數做一個很好的統計，就會發現在大部份的國家其實都分佈在這個迴歸曲線上，但有一些國家則是偏離這個迴歸曲線，可能在回歸曲線上面或是下面，有一些國家它的識字率很高，是無法從國家財富、都市化程度，及就學率還有師生比等這幾個因素來解釋的，這時候我們再來說也許是這些國家做的很好，超過大家的預期，那就來談談哪些國家做的很好？哪些國家超標最多呢？

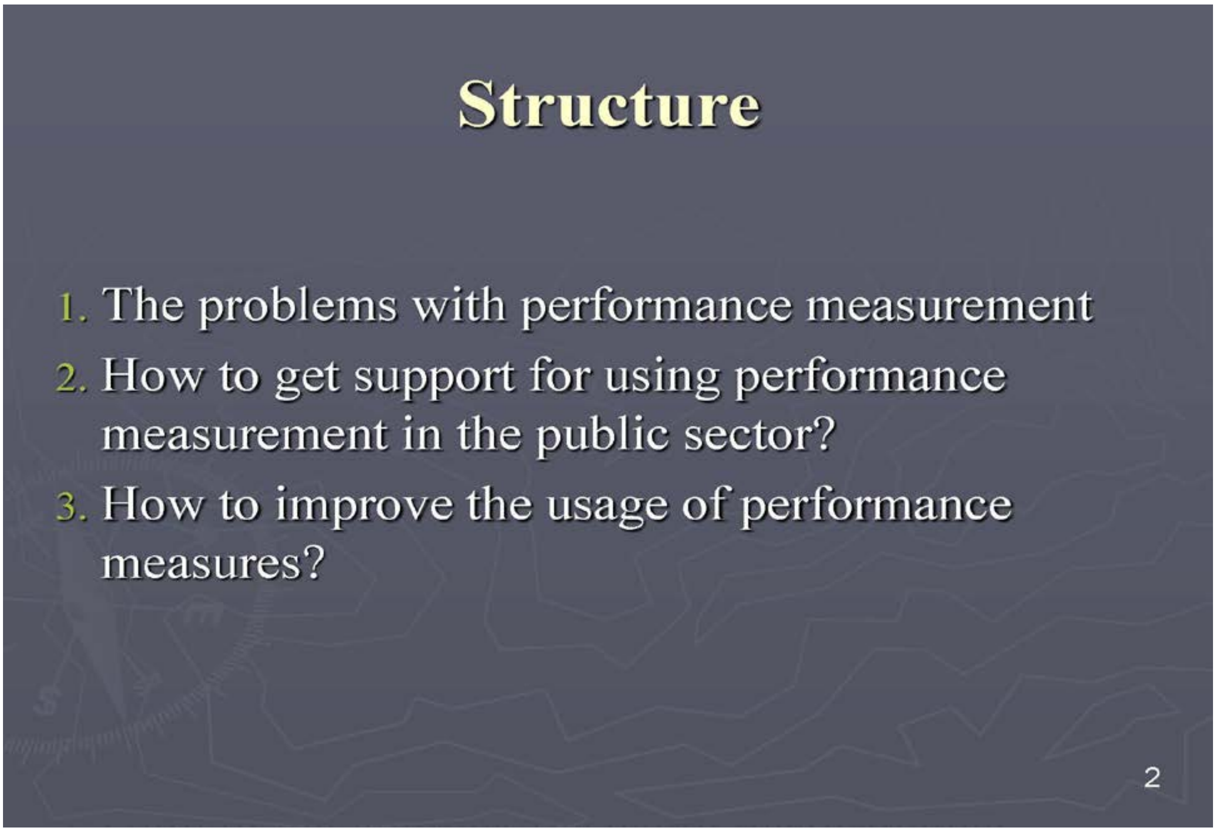
就芬蘭來說，它並沒有百分之百的識字率，但它的確是符合剛剛所說的因素，收入高、都市化程度高等等，現在我們要找一個國家，它的國家 GDP 是低的、人民的居住率是低的、都市化程度也低、就學率也低，還有師生比也不夠高，但它的識字率卻很高，大家覺得是哪個國家呢？是蒙古，蒙古幅員非常地大，GDP 也不高，但是識字率高，97.5% 的蒙古人都能夠識字，這是遠遠超出大家的預期，所以蒙古是做了什麼跟大家不一樣呢？

(如圖) 這個就是剛剛說到的識字率，蒙古排在上面，識字率 97.5% 這跟迴歸曲線差了很多，所以我做的研究是什麼呢？重點就是環境、情境，所以假設你的雇主、你的老闆說我們現在要做績效評估，請務必告訴你的老闆、你的長官，你要把環境、情境、相關條件放進去，這樣子你才能夠得到最正確的答案，知道自己的績效如何，而在公部門也是一樣，公部門當然有問題，有人會說績效評估不適用，公部門不能用績效評估，那我們怎麼做呢？我們要找出一個方法讓公部門更能夠接受績效評估，這個就是公共行政這個領域要做的，我們要找到解決這個領域的方法，這個也是大家所面臨的問題，謝謝大家。

How to use performance measures and to convince public official that they are indeed useful.

Michiel S. de Vries





“Should” give answers

- How well the organization does perform?
- To ensure that staff does the right things?
- In which programs should money and time be invested?
- To motivate staff to improve performance
- To convince others that the organization does perform
- To identify performance in order to reward them and to use them as examples of best practices
- What works, what doesn’t
- What needs to be done differently in order to improve performance?

Performance measurement
The expectations

It enables:

- evaluation
- control
- budgetting
- motivation
- promotion
- rewards
- learning
- improvement

1. The problem

5

Public Administration can help the public sector

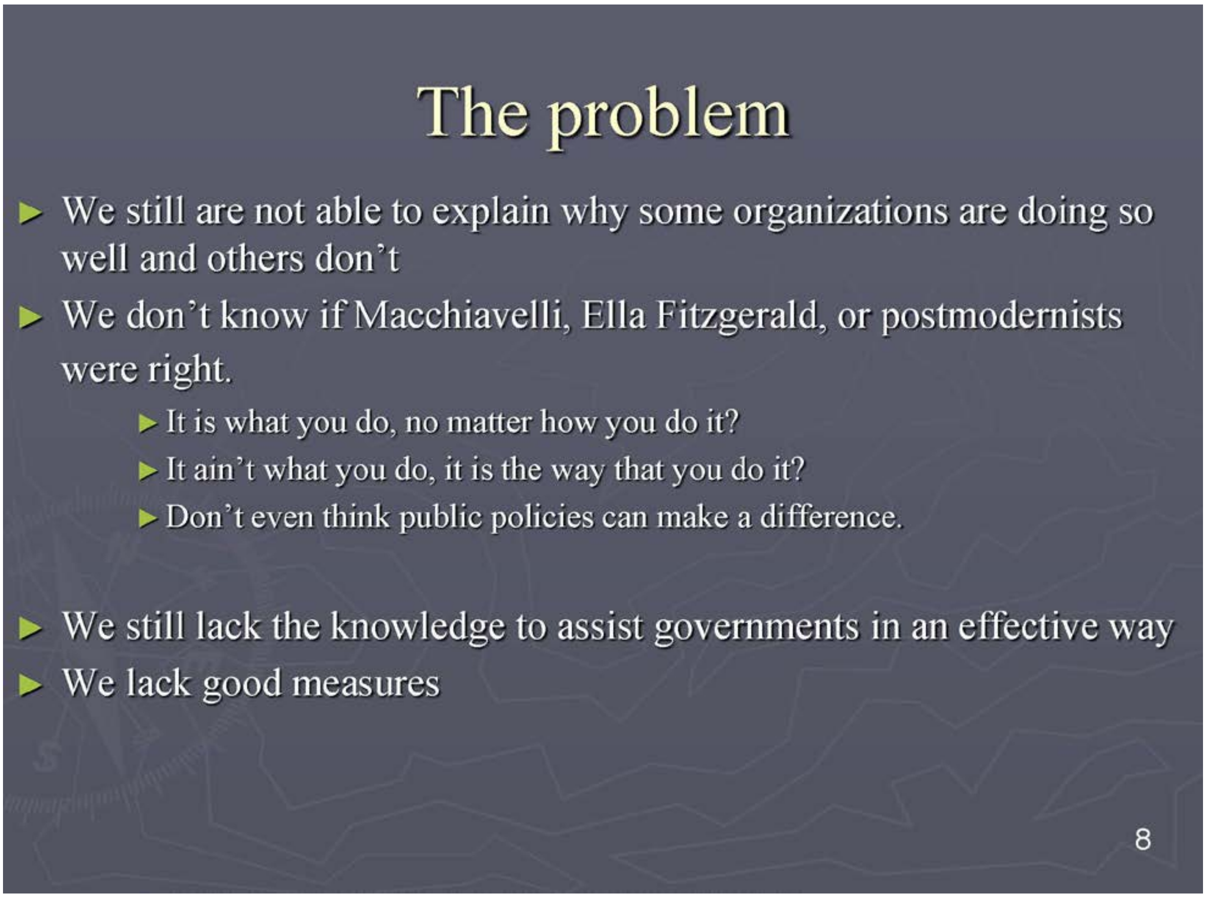
- To respond effectively to new developments
- To perform better

- **BUT**
- We have a long way to go

- I will argue that
  - Perhaps many organizations do not perform well
  - Perhaps practices change very slowly
  - But we in Public Administration lack the information to understand the problems, to explain and to improve
  - There is a way out by taking performance more seriously

6


My approach would be:
► To understand what determines Governments' performance
► To explain why some organizations succeed and others don't
► To assist governments using the knowledge we possess
This implies
► Good measures
► Good analyses
► Convince public officials that this is in their own interest
7

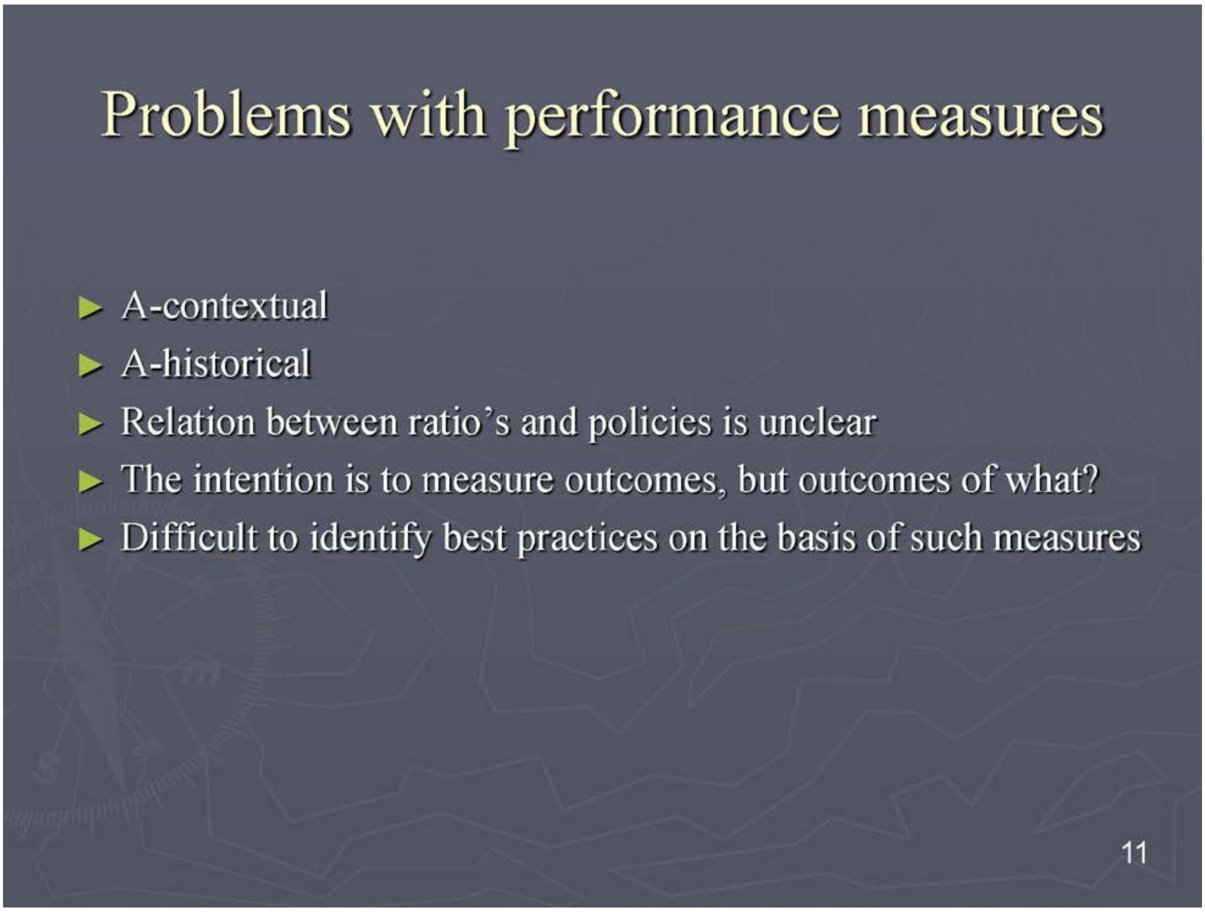

The problem
► We still are not able to explain why some organizations are doing so well and others don't
► We don't know if Macchiavelli, Ella Fitzgerald, or postmodernists were right.
► It is what you do, no matter how you do it?
► It ain't what you do, it is the way that you do it?
► Don't even think public policies can make a difference.
► We still lack the knowledge to assist governments in an effective way
► We lack good measures
8

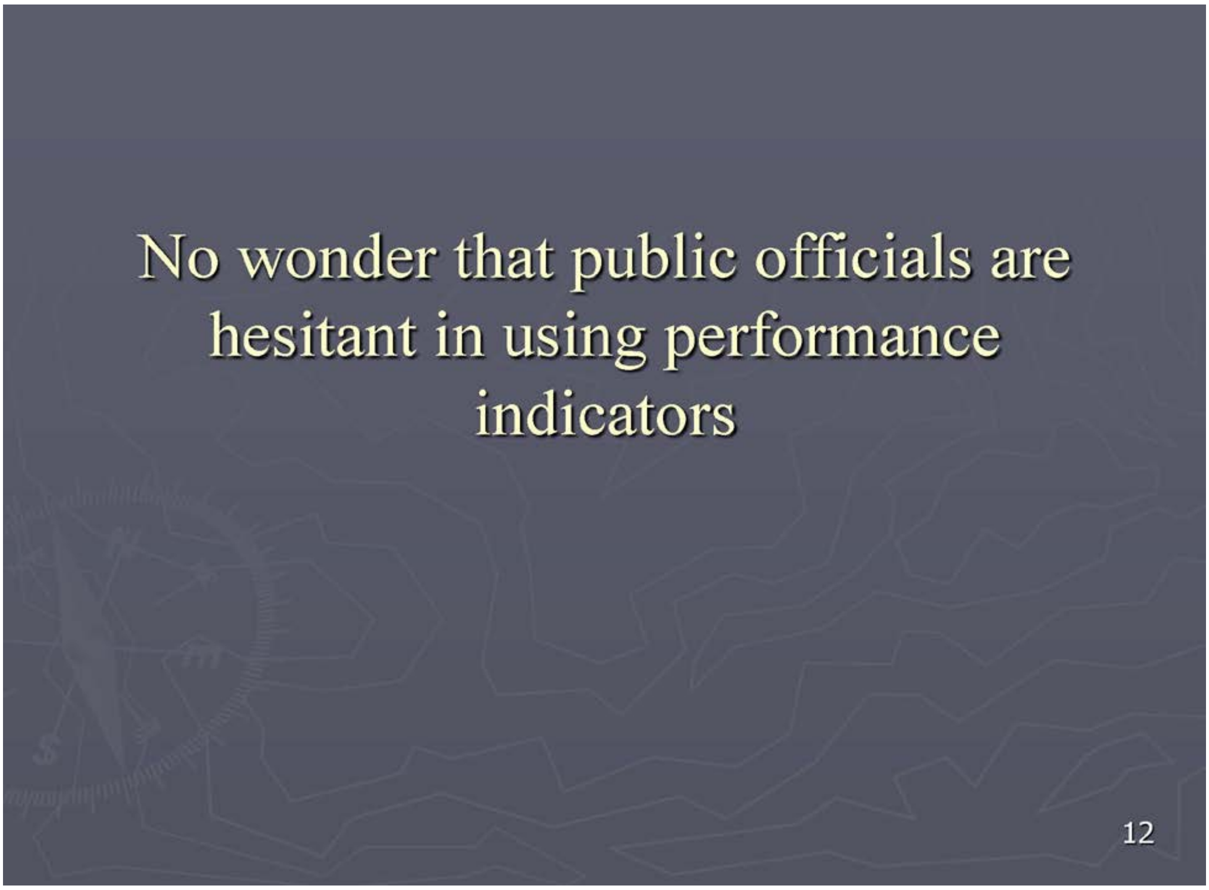
Problem in performance measures

- Often too simple
  - How much
  - How often
  - What percentage

What is and is not measured

Is measured	*Is not measured*
Input	Proces
Output	Outcomes


Problems with performance measures
A-contextual
A-historical
Relation between ratio's and policies is unclear
The intention is to measure outcomes, but outcomes of what?
Difficult to identify best practices on the basis of such measures
11


No wonder that public officials are hesitant in using performance indicators
12

2. Is performance measurement applicable in the public sector?

Performance measurement in the public sector is controversial

Performance measurement may be dysfunctional, have unintended, and even adverse consequences for the quality and productivity and can contribute to a reduced credibility of government among the service recipients of that government (Halachmi)

My research

How do public officials themselves feel about the opportunities of the application of public management in view of the measurability of their performance and how can the differences in their views on this be explained?

15

Database

- A large-scale survey research conducted under the auspices of the Dutch Ministry of the Interior and Kingdom Relations (MWM2, 2010).
- 26,876 civil servants participated in this study, which was held at the end of 2010.
- Officials were asked, for instance, to indicate to what extent they think performance agreements about their work are well measurable.

16

Theoretical expectations regarding the effect of factors on the opinion about the applicability performance according to three models			
	Model 1. The difference between the public and private sector is crucial	**Model 2. Substantive, rational arguments dominate**	**Model 3. Public choice model**
Expected influence of:			
1. Sector: Public versus private	Negative *	none **	Positive ***
1. Context: politically steered or not	None	Negative	Positive
1. Function Management of policy	None	Positive	None
1. Function: Complex or not	None	Negative	Positive
1. Experience with performance measurement	None	None	Positive
1. Quality of functioning by colleagues	None	Negative	Positive

* Within the model in which the specific nature of the public sector is central, a negative association between the type of sector (public versus private) and the views on the applicability of performance measurement is expected
**Within the model in which a technical rational perspective is central, it does not matter for the opinion about the applicability of performance measurement whether one refers to the public or the private sector as such.
*** From the public choice model in which personal and or organizational interests are central it may be expected that especially in the public sector, most proponents of performance measurement can be found.

17

Outcomes

My performance is well measurable (Public sector)

	Frequency	percentage
Completely disagree	1100	4,1%
Disagree	4153	15,5%
Neither agree nor disagree	5813	21,6%
Agree	10445	38,9%
Completely agree	2931	10,9%
Don't know	2334	9,1%
Total	26876	100%

On average respondents in the private sector and public sector do not differ so much when it comes to their views on the applicability of performance measurement (47% versus 53% in favor). Employees in both sectors are divided on the proposition whether their results are well measurable.

18

Explaining the attitudes towards performance measurement

	Sig.	Exp(B) Odds ratio
Experience with performance measurement	0,000	3,653
At least one performance review last year	0,000	1,541
Complexity of work	0,000	1,379
Political context	0,000	1,209
Quality of work by colleagues	0,000	1,160
Nature of work: management	0,002	1,094
Constant	0,000	0,252

Presented are the outcomes of the logistic regression
Nagelkerke $R^2 = 0,165$, N = 23890.

19

Interpretation

- Public officials are more positive about performance measurement if they
  - Have experience in concrete work arrangements
  - Have a more complex task,
  - Work in a political context,
  - Are positive about the performance of their colleagues,
  - Have a managerial instead of a policy related position

20

Theoretical expectations regarding the effect of factors on the opinion about the applicability performance according to three models			
	Model 1. The difference between the public and private sector is crucial	**Model 2. Substantive, rational arguments dominate**	**Model 3. Public choice model**
Expected influence of:			
1. Sector: Public versus private	Negative *	none **	Positive
1. Context: politically steered or not	None	Negative	Positive
1. Function Management of policy	None	Positive	None
1. Function: Complex or not	None	Negati	Positive
1. Experience with performance measurement	None	None	Positive
1. Quality of functioning by colleagues	None	Negati	Positive

* Within the model in which the specific nature of the public sector is central, a negative association between the type of sector (public versus private) and the views on the applicability of performance measurement is expected
**Within the model in which a technical rational perspective is central, it does not matter for the opinion about the applicability of performance measurement whether one refers to the public or the private sector as such.
*** From the public choice model in which personal and or organizational interests are central it may be expected that especially in the public sector, most proponents of performance measurement can be found.

Conclusions of this part

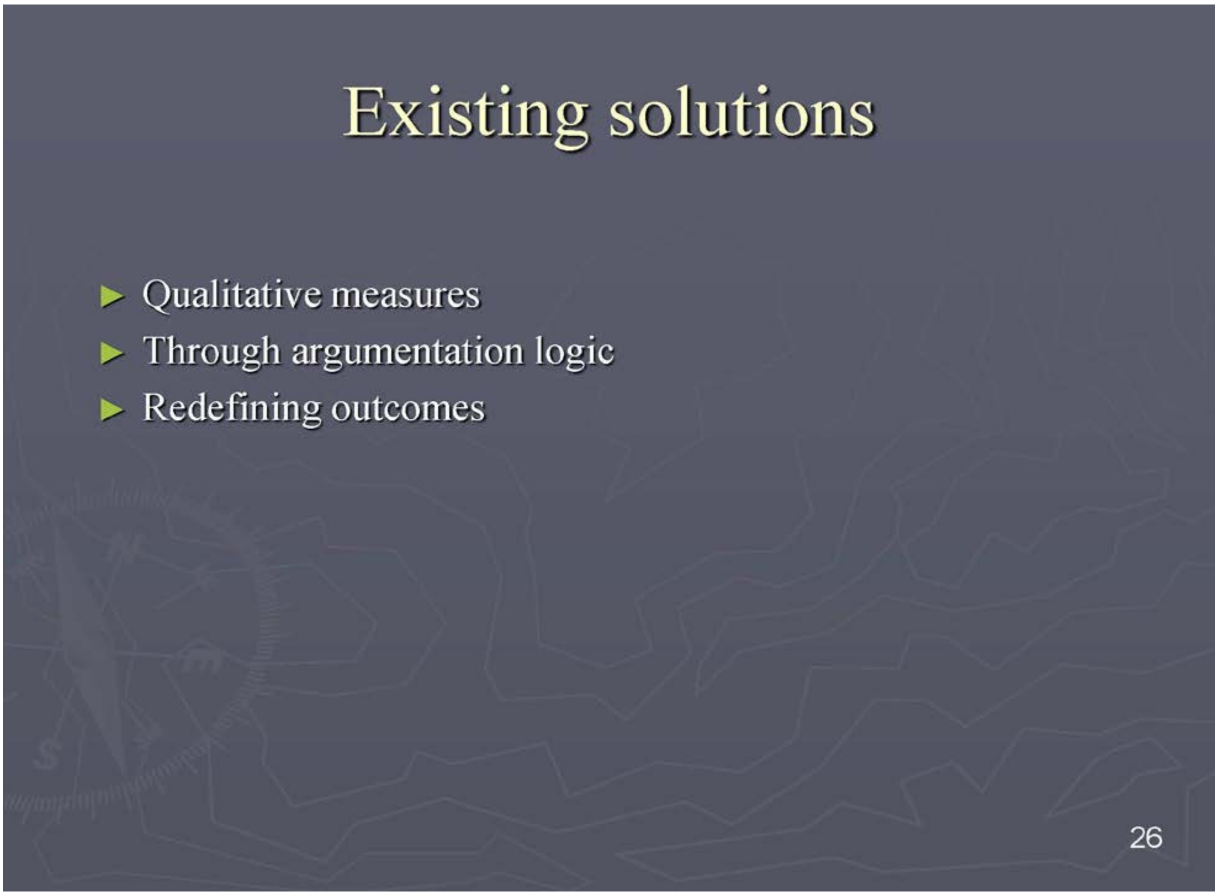
- The outcomes are especially in line with the public choice model.
- Performance measurement as a protection of the public sector.
- The instrument could be very useful to reverse the negative image that most administrations in municipalities, provinces and water boards experience, into a more positive image.

22


But still, we do need valid and reliable performance indicators
23


3. Improving performance measurement
24


Measuring performance
► We do measure inputs and outputs
► We hardly measure process and outcomes
► In terms of Christopher Hood we should measure
▪ Sigma values: → Goal-achievement
▪ Theta values: → Integrity
▪ Lambda values: → Robustness and sustainability
► However, only Sigma values are measured
25


Existing solutions
► Qualitative measures
► Through argumentation logic
► Redefining outcomes
26

Qualitative performance measures

- Surveys
  - "Are you happy?" or
  - "Do you feel safe on the streets?"
- Visitation committees
- Inspections
- Audits (Mostly based on self evaluations)
- Awarding prizes and recognition of best practices

27

Through argumentation logic

- We measured the output and this implies given the following argumentation that it is self-evident that we achieved the outcomes .....

28

Redefinition of outcomes

- Put outputs and outcomes under the denominator of results
- See output as organizational goals, outcomes as effects to achieve societal goals
- See outputs as single indices, outcomes as a composite index
- Define outcomes as outputs at a more remote date in time
- Outputs are measured by absolute levels and the outcomes as changes
- See outcomes as measurements on which the parties can agree that they are outcomes

29

An example is the WGI Worldwide Governance Index

- Compiled by the World Bank
- To rank countries according to good governance
- Six indicators
- Measures from 1996 onwards
- Numerous files aggregated into composite scores
- Each country gets ranked
- Most of the data are based on opinions

30

Problem with performance measures is more general

In general performance measures are criticized for being:

- A-contextual
- A-historical
- Relation between ratio's and policies is unclear
- The impossibility to measure outcomes
- Difficulty to identify best practices on the basis of such measures

31

Alternatives

- Data gathering: The NEPAD method. Peer review by neighbour countries
- Data analysis: A five step procedure taking the context into account

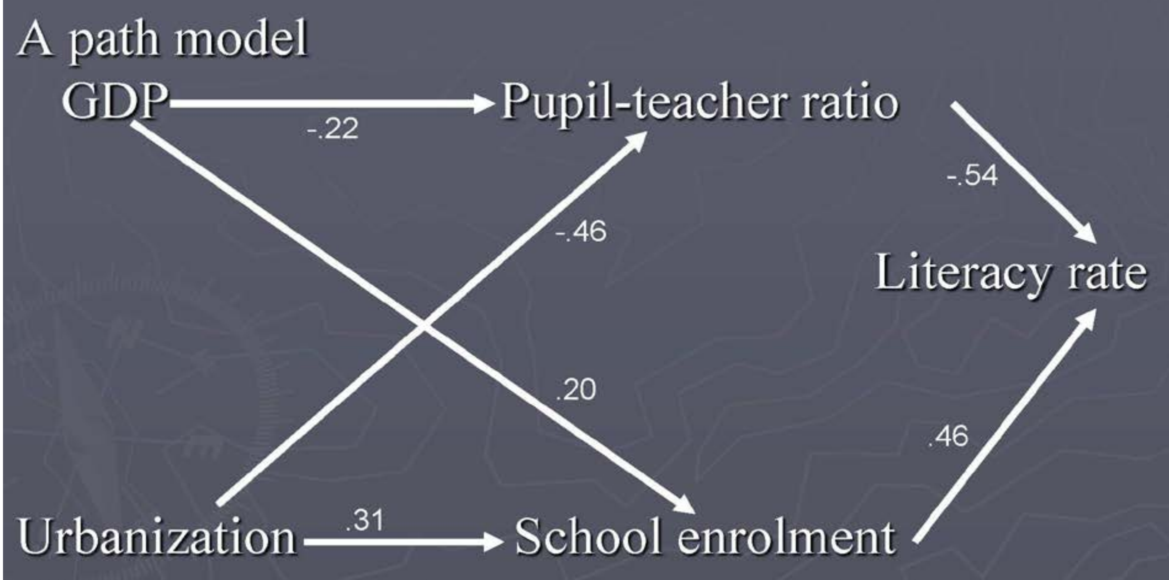
32

Data analysis: a five step procedure

- The construction of a basic causal model relating the policy problem to its causes found in the context.
- The testing of this model using statistical data and multivariate analysis
- The addition of policy outputs to the model
- Benchmarking based on the residues.
- Doing a content analysis on the policies of those countries in which the residue is largest

33

Application to literacy rates



34

Best performing countries given the residues

Table 3. *Best performing countries given the residuals*

Cambodia	74,97	13,14
Philippines	93,19	14,22
Nigeria	70,00	15,20
Namibia	87,24	19,09
Mongolia	97,46	19,46
Top 5 based on regression	Real literacy rates	Residue Real – predicted

35

Content analysis

- At the beginning of 20th century Mongolia had a literacy rate below 10 percent. Now it is up to 97%.
- Mongolian government used clear slogans that were communicated widely among the population and were followed up with financial support.
  - *Master literacy first, and then teach others*
  - *Summer literacy Campaign*
  - *Eye of Wisdom*

Results:

- Every child, even those from nomadic families in remote areas, could go to school.
- Every school had a full staff complement
- Student-teacher ratios were among the lowest in the world.

36

About this method

- Some scholars have criticized the method
- It is hardly used in Public Administration

- My conclusion is that:

The five step model using performance measures seems adequate to measure performance and to identify best practices

IF: the analysis is based on principles of:

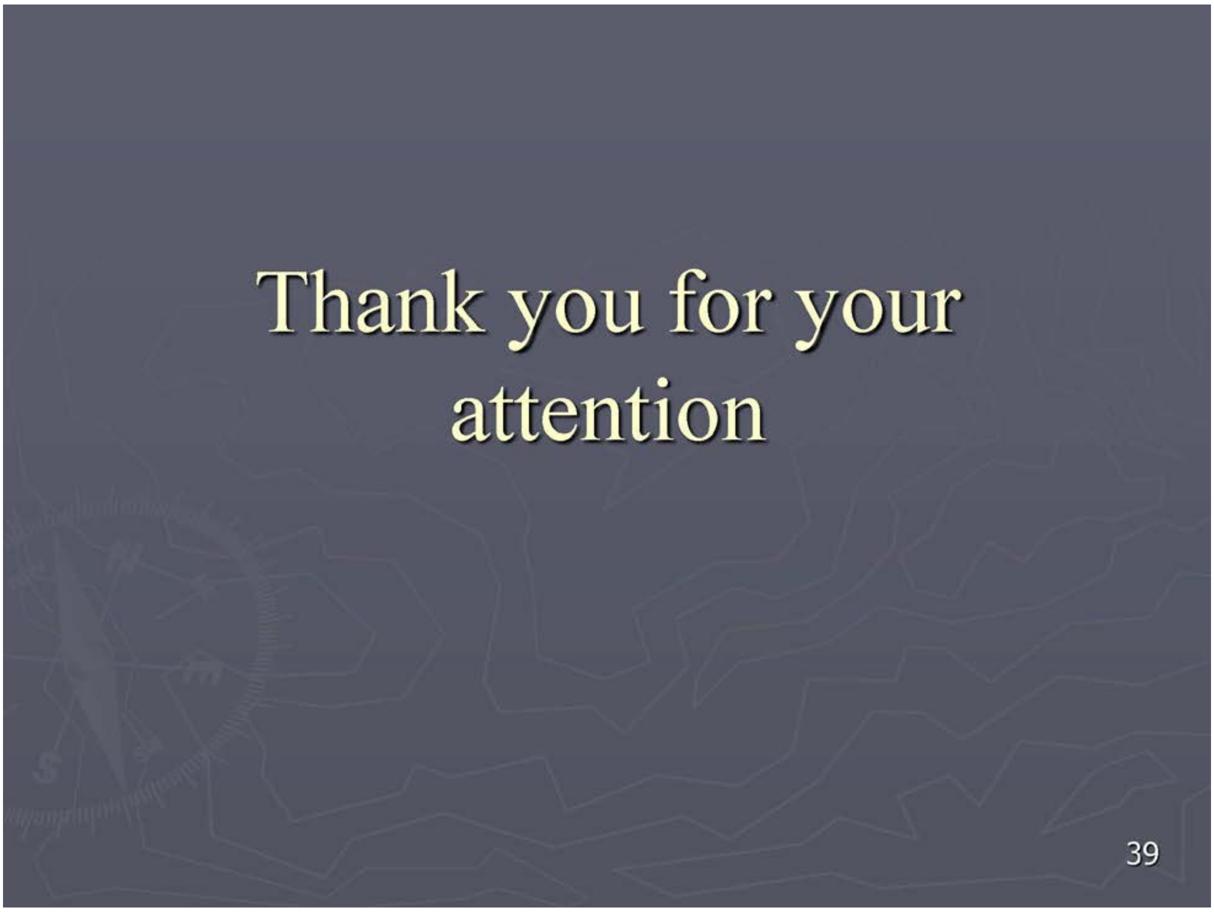
- Economy
- Theory
- Prudence

37

Conclusions

- Performance measurement is not as easy as it looks
  - There are a lot of expectations
  - But there is also a lot of criticism on its usage
- The public sector is in need of good performance measurement
  - According to officials themselves especially to improve their poor image
- But then we should invest a little more in the measurement of performance.
  - The five step procedure can be helpful in this regard.

38


Thank you for your attention
39


Appendices
40

WGI
Worldwide Governance Index

- Compiled by the World Bank
- To rank countries according to good governance
- Six indicators
- Measures from 1996 onwards
- Numerous files aggregated into composite scores
- Each country gets ranked
- Most of the data are based on opinions

41

Six indicators

- Voice and Accountability
- Political Stability and Absence of Violence
- Government Effectiveness
- Regulatory Quality
- Rule of Law
- Control of Corruption

42

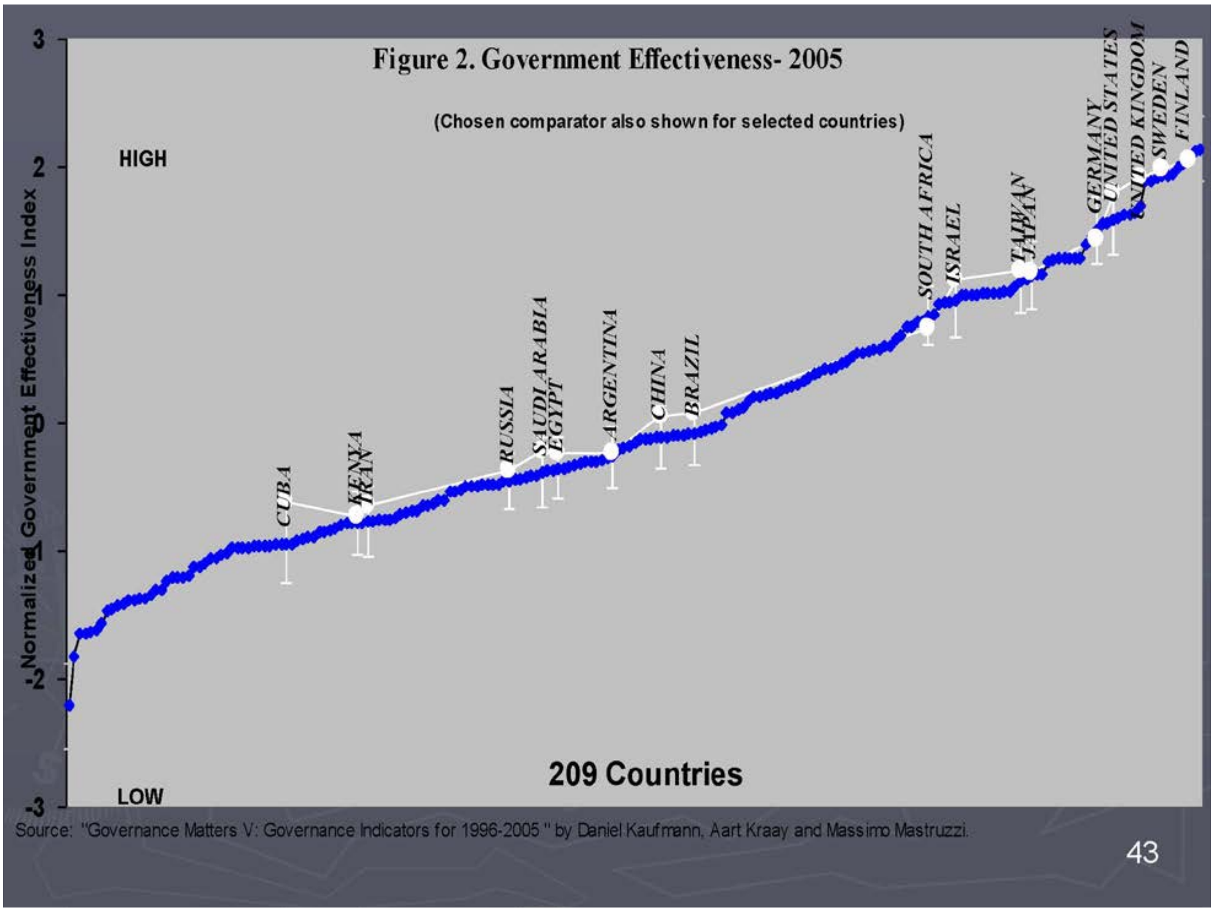
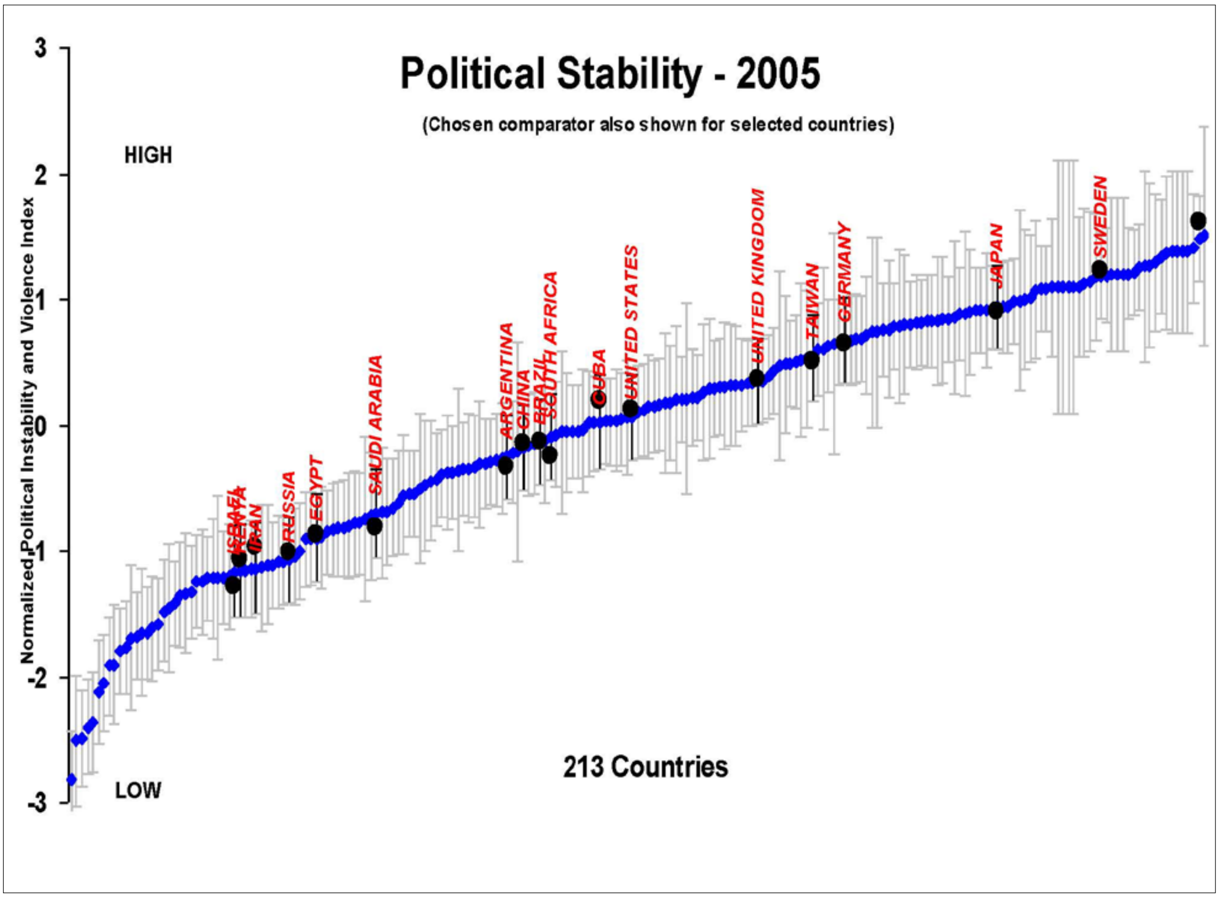
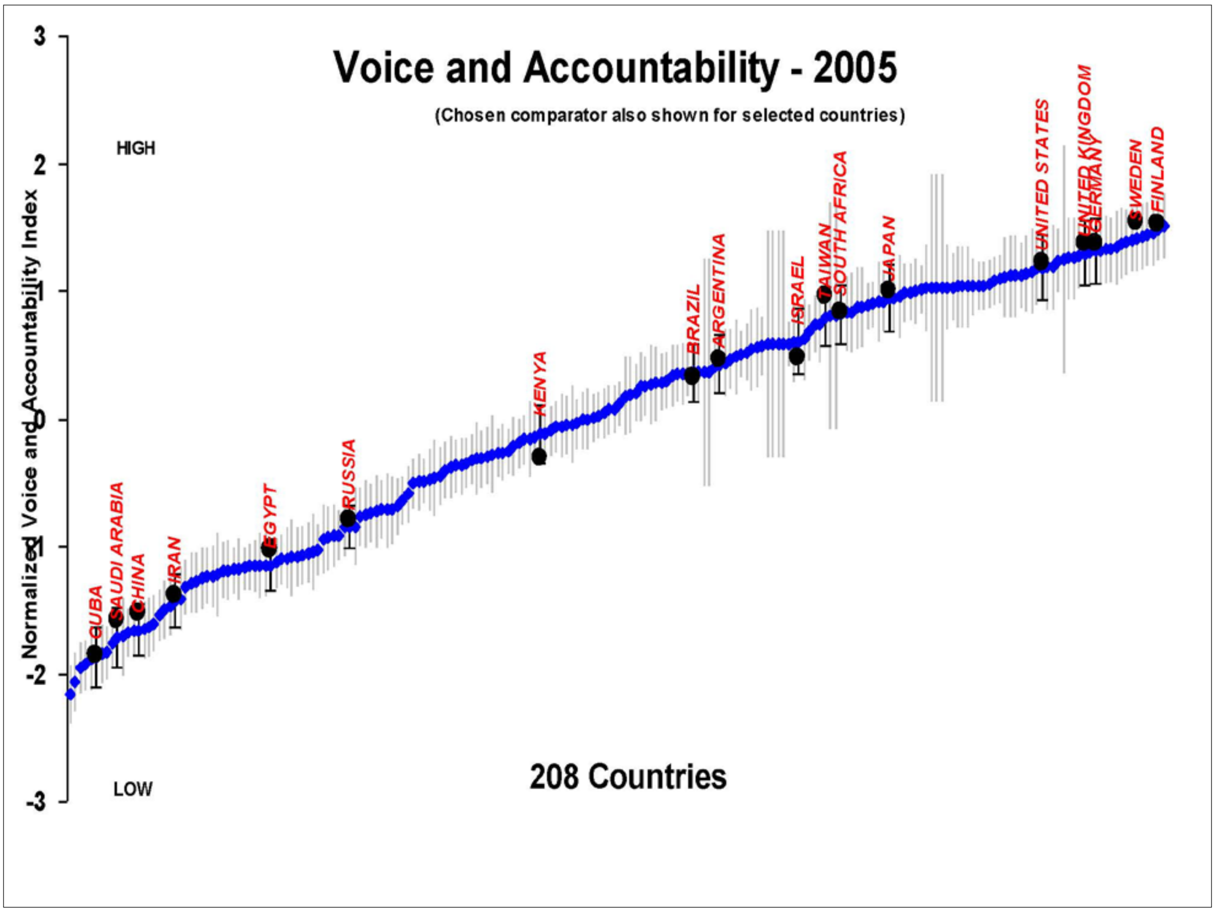
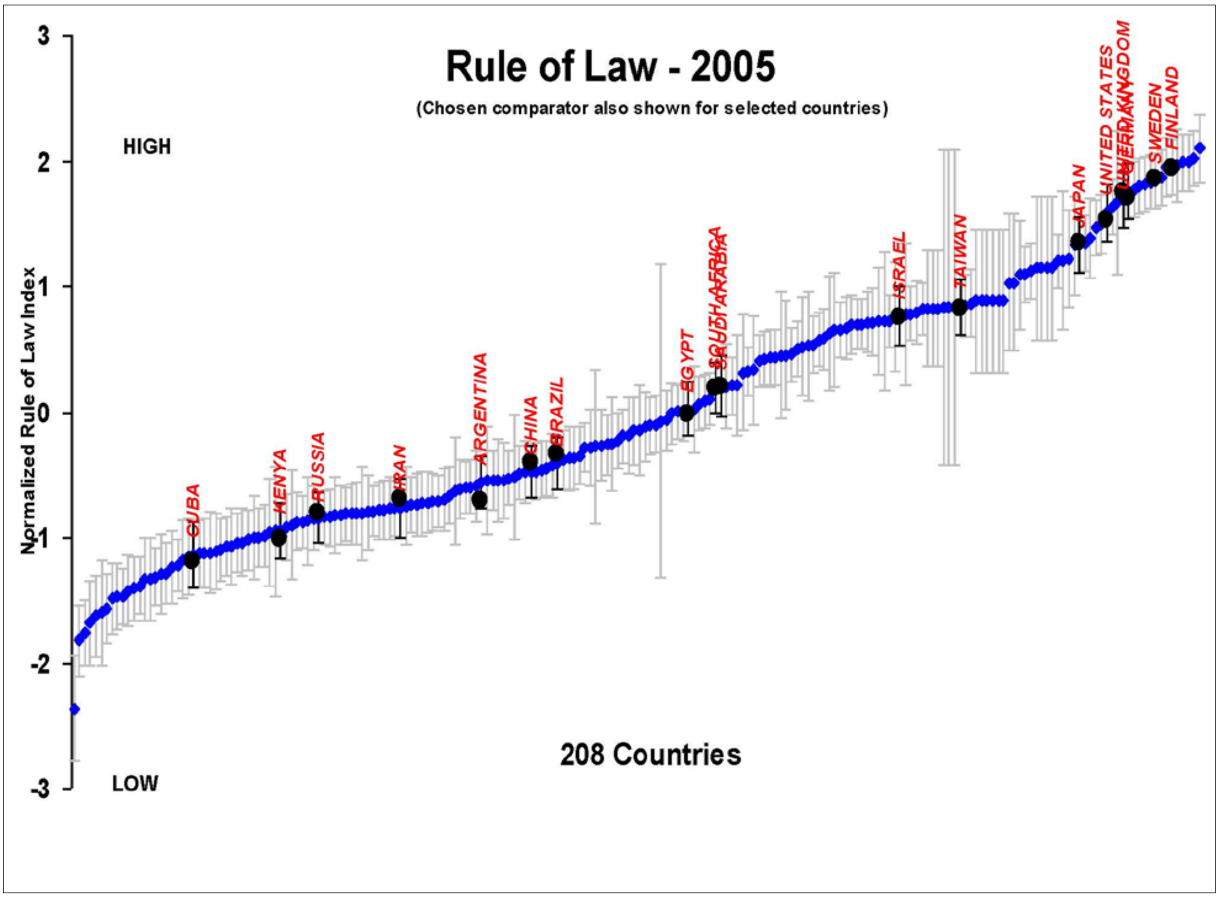
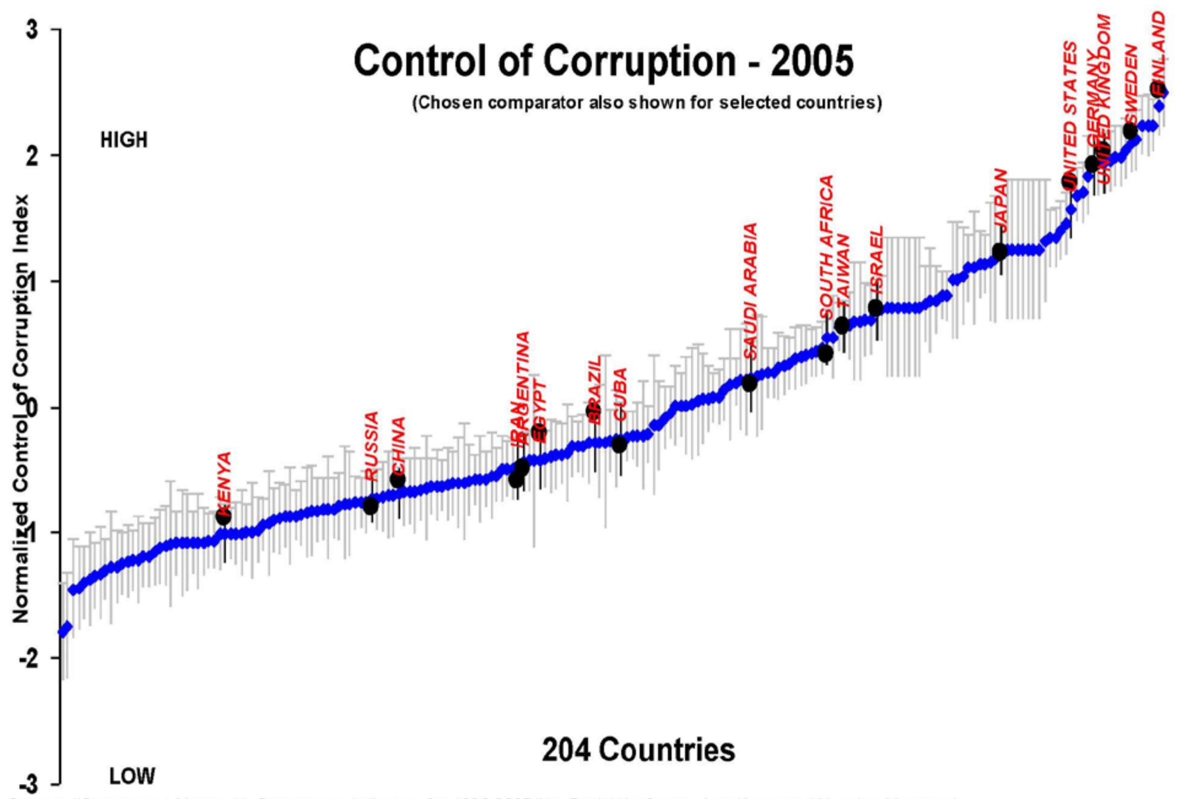


Figure 2. Government Effectiveness- 2005
(Chosen comparator also shown for selected countries)
HIGH
LOW
209 Countries
Normalized Government Effectiveness Index
3
2
1
0
-1
-2
-3
CUBA
KENYA
IRAN
RUSSIA
SAUDI ARABIA
EGYPT
ARGENTINA
CHINA
BRAZIL
SOUTH AFRICA
ISRAEL
TAIWAN
JAPAN
GERMANY
UNITED STATES
UNITED KINGDOM
SWEDEN
FINLAND
Source: "Governance Matters V: Governance Indicators for 1996-2005" by Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi.
43

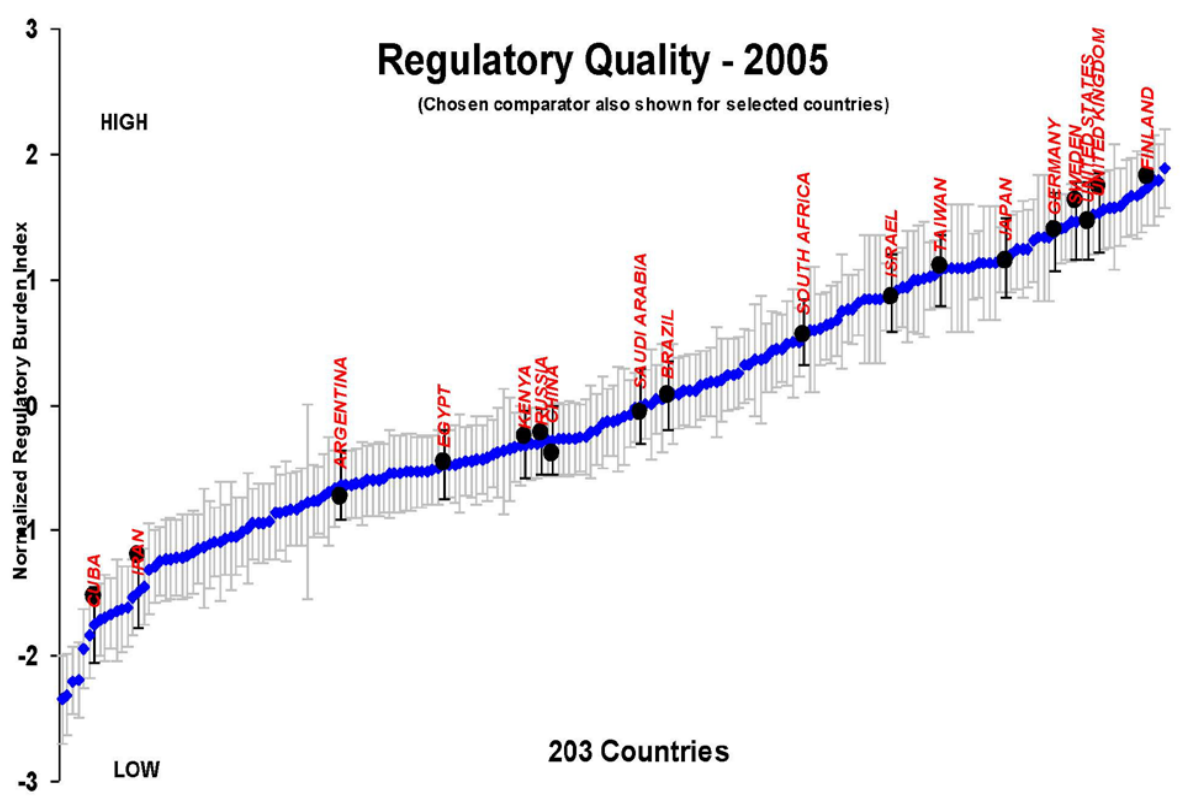

Political Stability - 2005
(Chosen comparator also shown for selected countries)
HIGH
LOW
213 Countries
Normalized Political Instability and Violence Index
3
2
1
0
-1
-2
-3
ISRAEL
KENYA
IRAN
RUSSIA
EGYPT
SAUDI ARABIA
ARGENTINA
CHINA
BRAZIL
SOUTH AFRICA
CUBA
UNITED STATES
UNITED KINGDOM
TAIWAN
GERMANY
JAPAN
SWEDEN


Voice and Accountability - 2005
(Chosen comparator also shown for selected countries)
HIGH
LOW
Normalized Voice and Accountability Index
208 Countries
CUBA
SAUDI ARABIA
CHINA
IRAN
EGYPT
RUSSIA
KENYA
BRAZIL
ARGENTINA
ISRAEL
TAIWAN
SOUTH AFRICA
JAPAN
UNITED STATES
UNITED KINGDOM
GERMANY
SWEDEN
FINLAND


Rule of Law - 2005
(Chosen comparator also shown for selected countries)
HIGH
LOW
Normalized Rule of Law Index
208 Countries
CUBA
KENYA
RUSSIA
IRAN
ARGENTINA
CHINA
BRAZIL
EGYPT
SOUTH AFRICA
SAUDI ARABIA
ISRAEL
TAIWAN
JAPAN
UNITED STATES
UNITED KINGDOM
GERMANY
SWEDEN
FINLAND



Source: "Governance Matters V: Governance Indicators for 1996-2005" by Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi.



Source: "Governance Matters V Governance Indicators for 1996-2005" by Daniel Kaufmann, Aart Kraay and Massimo Mastruzzi.

Problems in data collection

Based on

- **perceptions.**
- **Other databases**
  (Asian Development Bank, UN, Transparency international)
- **Intrapolation**
  (f.i data absent for Africa regarding corruption, but present for government effectiveness. Look how corruption and effectiveness correlate in Asia and insert data on corruption for Africa based on that regression analysis)

Criticism

- Often only 4 or 5 experts
- The database, consisting mainly of perception-based-cross-country measures of governance, would be flawed and that halo effects are present in which the wealth of countries and political ideology result in a biased ranking of countries on this governance scale. (Kurz and Shrank,2006)
- All these databases refer to one another
- A country can do nothing to improve its own governance in this respect.
- It is assumed that because effectiveness and corruption correlate in Asia they must be related in Africa

49

Problems in data analysis

Method

- All indicators correlating highly must refer to the same dimension of governance

Criticism

- The database is a complex, a-theoretical and yet poorly articulated hypothesis for which no evidence has been advanced. (Thomas, 2006)
- The stability that was found among nations may well be an artifact of the method used.
- Aggregating a multitude of indicators in one performance measure automatically results in stability

50

For instance: Government effectiveness

Is measured by subjective ratings of the:

- administrative and technical skills of the country's civil service
- efficiency of the country's national bureaucracies overall.
- efficiency of the country's local-level government bureaucracies overall.
- effectiveness of coordination between the central government and local-level government organizations.
- state's ability to formulate and implement national policy initiatives.
- state's effectiveness at collecting taxes or other forms of government revenue.
- Does the central government produce a national budget in a timely manner?
- Do local governments produce budgets in a timely manner?
- state's ability to monitor socioeconomic trends
- state's ability to create, deliver, and maintain vital national infrastructure.
- state's ability to respond effectively to domestic economic problems.
- state's ability to respond effectively to natural disasters.

51

But also by:

- Competence of public sector personnel
- Quality of general infrastructure
- Quality of public schools
- Time spent by senior management dealing with government officials
- Trust in Police

52

Concluding

These data

- lack validity,
- May well have a Western bias
- Result in the artificial conclusion of stability of bad governance and the near impossibility to improve
- Confuse theoretical terms like effectiveness, efficiency, and ability
- Blur explaining factors, indicators and impacts of governance

53

Reply from Kaufman et al

- Given the advanced statistical methods used,
- Given the multitude of independent sources,
- Given the high correlations between indicators,

(ACCORDING TO THEM)
THE CRITICISM IS HOSTILE,
BESIDE THE POINT
AND EXAGGERATED

- According to them: The reliability of the measures must be high and therefore the validity is high

54

My suggestion

- The work done by the World Bank should be done and is important
  - For distinguishing between good and bad governance
  - For understanding the problems of governance
  - For explaining the problems behind governance
  - For targeting one's solution
  - And for improving the state of states
- But we run the risk that the comparative approach ends the same way as the comparative public administration research movement ended in the 1960s and 1970s, namely in disappointment
- That is because there are severe problems and one should not use these data without caution

55

How to use performance measures and to convince public officials that they are indeed useful

Michiel S de Vries[1]

1.Introduction

The use of performancemeasurement andperformance managementinthe public sector remainsa controversialissue.Performance managementis about achieving goals and measuring the extent to which these goalshave actually beenachieved. While the use of performance management in the private sectorhas increased (see Neely,etal,1997),and there are some staunch advocates for its use(seeBroadnax,etal, 2001; Townley,etal, 2003; Behn, 2004) based on the identified benefits ofencouraging individuals to achieve pre-setgoals(Lohman, 1999), there is still some doubt as to whether themeasures are a valid, accurate reflection of the contentsof the work.In other words, do they measurewhat they'reintendedtomeasure?(See Steers, 1975; Kanter& Brinkerhoff, 1981; Kaplan and Norton 1992; Judge 1994; Guthrie, 1994; Au 1996; Anspach 1991; Guthrie and Parker, 1998; Kluvers, 1998; Young, Oakes and Preston, 1998; Kouzmin, et al, 1999; Pollitt and Bouckaert, 2000; Spigelman, 2001; Bovaird&Löffler, 2003; Bevan & Hood, 2006; De Vries, 2010; Halachmi, 2011)

Scholars havearguedthatthe complexityof public administrationwith its various stakeholders prevents the use of simple performance measures since they would not reflectthe nature ofpublic services.Some suggested using differentparameters in conjunction with the performance measures (Boschken, 1994, p. 312). Based on their health care research, WernerandAsch(2007)go further andeven conclude that: “performance measurement gets in the way of delivering good care, because it risks diverting attention from other more important but unmeasured aspects of care, provide(s) no priority for following guidelines likely to yield a large clinical benefit compared with guidelines likely to yield at best a small clinical benefit”, and “may crowd out quality at the level of the patient that is equally important but that cannot be easily measured. Hence,they reducethe quality thatis most important for the patient butcannot be easilymeasured” (cf.Halachmi, 2011, p. 25).According toHalachmi(2011,p. 37), there is no guarantee that theobjectivesof performance measurements can be achieved.Performance measures maybedysfunctional, have unintended, and even adverse consequencesfor quality andproductivityand can contributeto reducinggovernment credibility among theservice recipients.

I will address these issues, not by taking sides, or by telling you, to use or don’t use performance measurement and management, but by going the way Public Administration usually

1.Prof. dr. M.S. de Vries is full professor and chair at the department of Public Administration, Radboud University Nijmegen, the Netherlands. E-mail: m.devries@fm.ru.nl

proceeds. That is, we see that something is desirable, in practice we see that this desirable thing is not working the way it is supposed to work and therefore we define it as a problem to be solved. The next step is to investigate how to get reality closer to the ideals? This results in recommendations for improvement to reduce the problems.

In Public Administration we are pragmatic and critical, but not cynical. If one sees something problematic, one can just criticize and leave it with that. That is the cynical approach. In Public Administration we do take the first step, that ism to critically assess the problem, but we proceed differently. We see something wrong, faulty, or problematic and then start thinking. How to resolve the problem? This implies investigating the causes of the problem. To take the removal of these causes as our goals and find instruments to achieve these goals.

How does this apply to performance measurement?

I see two main problems:

1.There is little support for the application of performance measurement and management in the public sector.

2.The performance measures themselves are problematic.

So what to do? Let’s look at the problems, investigate what is causing them and think about possibilities to resolve them. That is what is done below.

2.The absence of support for performance measures in the public sector

Why is that the case? Why are so many public administrators hesitant, if not outright opposed to performance measurement? If we investigate the objectives of performance measurement they seem to be only advantageous. What objectives did I find in the literature?

Performance measurement would enable evaluation, control, budgeting, motivation, promotion, rewards, learning and improvement. It should give answers to pressing questions such as: How to ensure that staff does the right things? How to motivate staff to improve performance? What works, what doesn’t and, what needs to be done differently in order to improve performance?

When you put all these objectives found in the scholarly literature together you see immediately what is amiss. All these goals presuppose that things can be improved, or even that at the moment things don’t go as they are supposed to go. It is a negative assumption underlying performance measurement and management, namely that the public sector is not working efficiently and effectively and that by measuringits performance you can support that negative view, pinpoint those problems and try to resolve them. In the extreme case performance measurement is a threat to the public sector and all those working for the public cause. My point is that you won’t get much support from public managers and administrators out of such a negative framing. I did some research in my home country, the Netherlands. The Ministry of Home Affairs conducts a survey amongst all public administrators every two years. In 2010 it resulted in a survey with a response of 26,876 public servants returning the questionnaire. They work in Central government, local government, provinces, the judicial sector, water boards, education and science, research institutes, university medical centers, defense and police. They were asked about the applicability of performance measurement to their work, the nature of their work - be it management or policy making, the complexity of their work, their experience with performance measurement and the quality of the functioning of oneself and colleagues.

Well now, I don’t know what your expectations are about the relationship between the applicability of performance measurement and these other factors, but I can tell you what the scholarly literature would expect.

Part of it, summarized here under the private versus public model, would expect that because of the complex nature of the public sector, performance measurement is not applicable at all. Au(1996)andAnspach(1991) argued that the performance ofan organization is asocial constructthatisparticularly difficult to measureinthe public sector.

A second stream, summarized here as the rational choice model, would predict thatthe attitudes towards the applicabilityof performance measurementand managementvary withinthe public sector. Ifstakeholdersbelieve that there are problems in the performance within the organization,they will see performance measurement as necessary and feasible. From this model, we expect a highersupport for the applicability of performance management among thoseofficials whoworkin a lesspolitical environment, whose functionis less complex, and especially among those officials holding managerial functions,who are dissatisfied aboutthe functioningof the organization.

The third model, summarized here as the public choice model, tries to explain the attitudes by the personal and organizational interests of the administrators. If performance measurement is in their personal or organizational interest they will favor it and otherwise oppose it. In thismodel,performance measurement andperformance management are judged for their contribution in improving the employees’ own position withinthe organization or to protect,defendand improve the position of the department within theorganization.From thismodel, one would expect to find a highpercentage of people to have positive attitudes regardingthe applicability of performance measures in cases where uncertaintyabout the functioning is great and the organization could profit from performance management. This implies that we would expect to find a higher percentage of proponents within the public sector agency that receives more criticism of itspoor performance than in the private sector.Evenmore proponents wouldbe foundamong officialswithcomplex functionsin apolitically steered organization wherethe criticism of bureaucracy is greatestand when officialshave positive experienceswithperformance management.Moreover,it can be expectedthatthose officialswho feel thatthey andtheir colleaguesfunction well are also more positive aboutthe applicabilityof performance measurementthanofficialswho are critical about the performancewithintheir organization.Inthe latter case,performance measurementis notin their interest.

Well now, in order not to bore you with all kinds of statistics, this is what the analysis showed in the end: 52% of all public sector employees favored performance measurement. Public officials are more positive about performance measurement if they have experience in concrete work arrangements, have a more complex task, work in a political context, are positive about the performance of their colleagues, and have a managerial instead of a policy related position. This is especially in line with the public choice model, that is, performance measurement as a protection of the public sector and in accordance with the interests of the organization. According to the outcomes of that investigation, the instrument could be very useful to reverse the negative image that most administrations in municipalities, provinces and water boards experience, into a more positive image.

It is one thing to tell public sector employees that it is useful to use performance measurement and management to address the inefficiencies and ineffectiveness of their work; that it can result in improvements; and that it is used to control and to monitor. This is what I would call the negative approach. It is a quite different thing to tell a positive story to the same people, namely that performance measurement can make it clear to the hostile outside world that the officials in the public sector do a difficult and complex job and is able to show that their performance is outstanding. The second way is much more productive and will result in much more support for using performance measurement within the public sector.

3.The problematic of performance measures

But, even so, critics will argue, performance measures in general do a lousy job. They do not provide a reliable and valid picture of the performance in the public sector. Performance measures do measure inputs and outputs, but hardly measure process and outcomes. Especially the problem that outcomes are hardly measured is very problematic. Furthermore, most performance indicators are a-contextual and a-historical. They result in awkward conclusions that if a problem increases, the performance of the part of public sector that addresses this problem, must have been decreasing. This neglects the fact that the development of societal problems can have all kinds of causes and the performance of the public sector is at best only one of those causes. Performance measures are too simple and do not do justice to the complexities involved in public sector work. They usually result in rankings that measure the extent of the policy problem, but they fail to reveal anything substantive about the quality of policies, and they are equally unhelpful in the search for best practices. Bovaird&Löffler stated that: "there are two key areas in which measurement is required: improvements in public policy outcomes; and implementation by all stakeholders of a set of principles and processes by means of which appropriate public policies will be designed and put into practice. (Bovaird&Löffler , 2003: 217). According to these scholars, there are important trends visible in the process of performance measurement, but "While during the first generation of public sector reforms in the 1980s and 1990s a lot of lip service was paid to the need to measure outcomes, not much action really took place on this front" (Ibid)

Outcome measures are often still directed at measuring input, process and output, and are unable to truly measure outcomes (Dijkstra, 2008). The problem of measuring outcomes persists because scholars measuring outcomes have difficulty relating outcomes to actual policies. This is all the more problematic because outcomes are also contingent, i.e. dependent, on the external circumstances while performance measures are mostly a-contextual. This contingency results in one of the main problems in outcome measurement, namely, how do we determine whether the change in the problem is caused or induced by the policy that intended to reduce or resolve the problem. It is difficult to answer whether it was the policy alone that was ineffective, because there are many factors in the specific context that could also have been the catalyst for the visible change.

Well now, it is one thing to criticize, which is justified in this case. It is quite another thing to become cynical. Again, when I see a problem I try to resolve it instead of just criticizing the situation. How to proceed?

I propose a five step procedure to improve the performance of performance measurement. The five steps involved are:

1. Construct a causal model that relates the policy problem to its causes found in the context in which the problem occurs, on the basis of existing theory (In order to make performance measurement contextual and less dependent on uncontrollable factors).

2. Test this model using statistical data and multivariate analysis over different organizations.

3. Add policy outputs for every organization to the regression model (In order to measure the added value (effectiveness) of the countries' policy controlling for the contextual factors. In this way, the effectiveness of policies can be measured)

4. Measure the residual (i.e. the difference between the expected extent of the problem and the real extent of the problem in all countries). The values of the residuals for each situation are used to determine those organizations which likely use a best practice.

5. Conduct content analysis into the policies in those circumstances in which the residual is largest

(in order to analyze what it is that makes the problem in this situation so much smaller than expected.)

An example may help to understand this five step procedure:

The variance in the number of crimes in municipalities in the Netherlands results in interesting rankings (see a/o Politie en Wetenschap, 2008). Based on such figures as collected on websites such as www.watdoetjegemeente.nl one is able to produce the following ranking for 2006.

Table 1. *Top 10 municipalities on number of crimes per 1000 capita in 2006*

Most unsafe municipalities		*Most safe municipalities*	
Utrecht [1]	47.5	Ferwerderadiel[2]	18.2
Amsterdam	143.6	Littenseradiel	19.1
Eindhoven	125.5	Leeuwarderadeel	19.6
Maastricht	121.1	Liesveld	19.8
Ouder-Amstel	120.1	Opsterland	21.0
Roermond	119.0	Ten Boer	21.4
Nijmegen	119.0	Graafstroom	21.9
Rotterdam	118.0	Menaldumadeel	22.0
Hertogenbosch	116.1	het Bildt	22.2
Arnhem	112.6	Ouderkerk	23.2

1. These are mostly large municipalities.
2. These cities are hardly known outside the Netherlands. This is because they belong to the smallest municipalities in the country.

The number of crimes per capita over all Dutch municipalities is a fairly normal distributed variable, somewhat skew with a mean of 55.3 and a median value of 51.9 crimes per 1000 inhabitants in which 95% of all municipalities score between 53 and 57.5. From recent scholarly research it is known that the degree of urbanization (Faris& Dunham, 1965; Traub& Little, 1999), poverty (Traub& Little, 1999), cultural diversity (De Beer &Schuyt, 2004) influence the crime rates. These contextual factors can hardly be tackled by public policies on security.

Table 2.Outcomes of the regression model on crimes in Dutch municipalities in 2006	Unstandardized Coefficients		Standardized Coefficients	t	Sig. [c]
	B [a]	Std. Error	Beta [b]	26,968	,000

Model 1	(Constant)	37,753	1,400		11,459	,000
	Children in special needs families	3,438	,300	,489	4,695	,000
	population 2006	,058	,012	,199	2,070	,039
	Percentage of immigrants	,370	,179	,080		
R^2 Model 1 = 0,40						
Model 2	(Constant)	67,742	2,889		23,451	,000
	Children in special needs families	4,046	,269	,576	15,060	,000
	population 2006	,011	,011	,182	4,879	,000
	Percentage of immigrants	,314	,157	,068	1,997	,046
	Percentage of crimes solved	-1,412	,123	-,376	-11,472	,000
R^2 model 2 = 0,56						

Dependent variable is crimes per 1000

a The unstandardized coefficient gives the effect of the factor in question on crimes per 1000 inhabitants

b These are the standardized regression coefficients

c. This is the level of significance of the regression coefficients.

The values for these factors is known in the Netherlands for each municipality (www.cbs.nl; www.waarstaatjegemeente.nl, www.watdoetjegemeente.nl) through indicators as the number of inhabitants, the percentage of children growing up in deprived families, the number of families living on social security, and the number of immigrants. The regression of these factors on the endogenous variable, results in an explanatory model that explains 41% of the variance in the crime rates at the municipal level (R^2= .41). The percentage of children living in deprived families is the most important explanatory factor, with a "slope" of 3.43 (β = .49, p<.000). Table 2 presents the outcome of this analysis in the upper model 1. All independent factors have the expected and significant effect on crime rates (N=442 municipalities, p<.05).

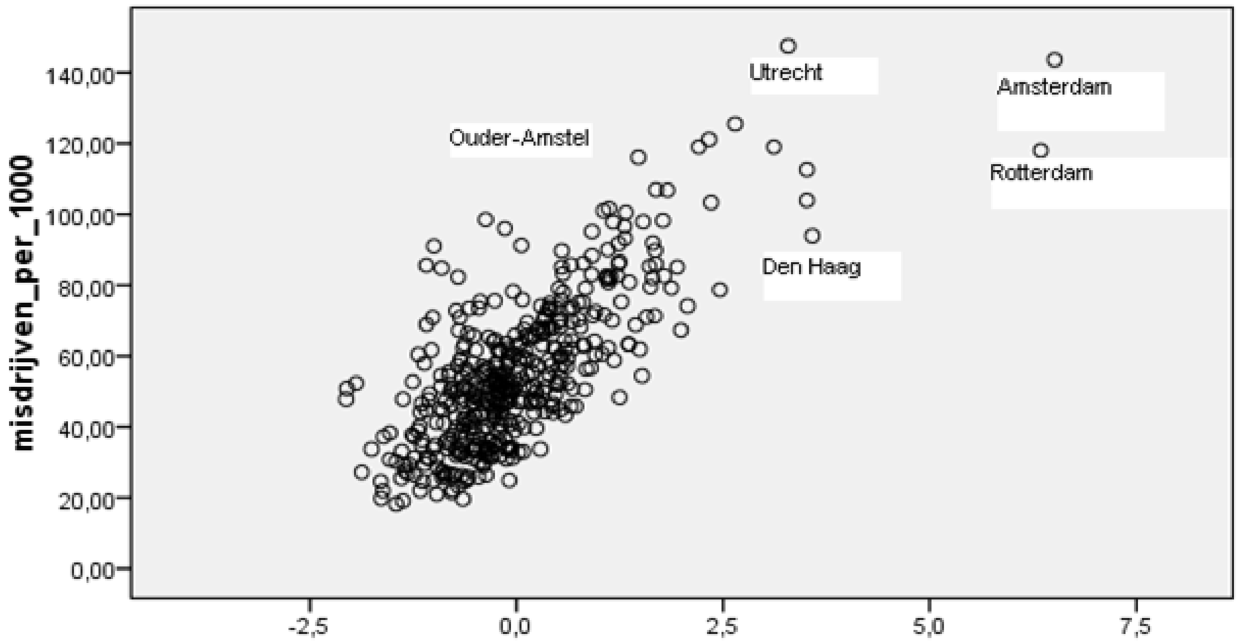
In model 2 we added the factor "risk of being caught" that is the number of solved crimes. The explained variance, R^2, increases to 55%. This implies an increase of the explanatory power by this one factor of 14%.

Searching for a best practice: The Rotterdam' coercion and urge' policies

The question arises whether there are municipalities that in all likelihood do conduct robust policies and where one sees a much lower crime rate than expected. Based on the regression analysis the expected value of the crime rate is related to the real crime rate. These are presented in the scatter plot given below. For some interesting municipalities the name is added to the scatter plot.

Figure 1.*Actual crime rates in 2006 and predicted values based on the regression analysis*

Actual crime rates



Predicted crime rates based on regression analysis(stand.)

The scatter plot shows a fine relation between predicted and real values of crime rates. However, there are some municipalities in which the difference between the two is rather large. Especially the position of the city of Rotterdam is outstanding. The actual crime rate in this city is almost 30% lower than one would expect on the basis of urbanization density, poverty, cultural diversity and the risk of being caught.

Did Rotterdam indeed conduct something like a best practice trying to cut the link between context and problem? An elaborate policy analysis is need to confirm this, but has to be disposed of because it would require a separate paper. Nonetheless the security policies in Rotterdam have changed significantly during the years under investigation. Dominant are phrases such as 'zero-tolerance', a tough approach against what are even in official policy documents called "shitheads", a tight policy on social housing in order to keep the needy out of certain neighborhoods (COS, 2008a) and a social security policy in which the principle "work first" is central (COS, 2008b). People, of whom it is expected that they abuse social security, or of whom it is known they cause nuisance, are forced to work at the garbage collection service. It is the so-called 'coercion and urge' approach. These aspects of security policies do distinguish Rotterdam from other municipalities.

The important thing, however, is not that Rotterdam is doing an outstanding job in reducing crime rates compared to other Dutch municipalities, but that performance measurement can be improved in such a way, as to make it more contextual and to find out what it is that causes a decrease or increase of societal problems and to pinpoint outstanding performance.

4.Conclusions

This presentation argued that critics are right in criticizing the dominant use of performance measurement and management in the public sector and that the support for performance measurement in the public sector is low. I argued that in Public Administration we are critical but must not cynical. We try to find solutions for the problems. That is our job.

In the case of lacking support for performance measurement within the public sector, the solution I proposed is not to stress the problems performance measurement is able to identify, but to emphasize the hard, thorough, complex job people in the public sector have to do and that they do perform. Performance measurement is for public administrators an adequate tool to convince the hostile outside world that people in the public sector work hard and perform well. Seen as an instrument against bureaucrat-bashing it might receive much more support within the public sector.

Secondly, the criticism towards the actual content of performance measures, as being a-contextual and a-historical, unable to relate outcomes to processes taking place in the public sector, is indeed justified, but a solution for this problem can be found or at least searched for also. The five-step procedure proposed above is a first attempt to resolve the problematic of performance measurement.

5.References

Anspach, R.R, Everyday methods for assessing organizational effectiveness, in: *Social Problems*; 1991, vol. 38, nr.1, p. 1-19.

Au, C.F, Rethinking Organizational Effectiveness: Theoretical and Methodological Issues in the Study of Organizational Effectiveness for Social Welfare Organizations, in: *Administration in social work*; 1996, vol. 20, nr. 4, p. 1-22.

Behn, R.D. (2003) Why Measure Performance? Different Purposes Require Different Measures, *Public Administration Review*, Vol. 63 No 5, pp. 586 - 606

Bevan, G. & Ch. Hood, Have targets improved performance in the English NHS?in: *British Medical Journal*; 2006, vol. 313, nr. 7538, p. 419-423.

Boschken, H.L, Organizational performance and multiple constituencies, P*ublic Administration Review*, 1994, vol. 54, nr.3, p. 308-312.

Bovaird, T, E. Löffler (2003). Evaluating the Quality of Public Governance: Indicators, Models and Methodology, in: *International Review of Administrative Sciences*;, vol. 69, nr. 3, p. 313-328.

Bovaird, T. and E. Löffler , (2003) Evaluating the Quality of Public Governance: Indicators, Models and Methodologies, *International Review of Administrative Sciences*, Vol. 69, No. 3, pp 313-328.

Broadnax, W.D. et. al. (2001)The Social Security Administration and Performance Management, in: D.W. Forsythe (ed.), *Quicker, Better, Cheaper? Managing Performance in American Government, Albany*, NY, November.

Cook, Th. Et al. (1995) Performance Measurement: Lessons Learned for Development Management, in: *World Development,* vol. 23, nr. 8, p. 1303-1315.

COS Centrum voorOnderzoek en Statistiek (2008a), *Evaluatie Wet Herhuisvesting Rotterdam*, Gemeente Rotterdam.

COS Centrum voorOnderzoek en Statistiek (2008b), *Werk Direct,* Gemeente Rotterdam.

De Beer P.T. & C.J.M. Schuyt (red) (2004) *Bijdragenaanwaarden en normen,* WRR, Amsterdam, Amsterdam UP.

De Castell, S., A. Luke & K. Egan (eds) (1986). *Literacy, society, and schooling: a reader* Cambridge University Press.

De Vries, M.S. & C.J. Lako (2009a) Prestatiemetingvooreffectief en robuustbeleid. *Bestuurswetenschappen,* forthcoming.

De Vries, M.S. & C.J. Lako (2009b) *Performance measurement and the search for best practices*.Paper at the IIAS conference. Helsinki. June.

De Vries, M.S. (2010)Performance measurement and the search for best practices, in: *International Review of Administrative Sciences*.Vol. 76. No 2, p 313-330.

Dijkstra, G. (2008) Outputmeting van overheidspersoneel: eensimpeleprestatie? *inBestuurswetenschappen.*No. 1. pp. 54-68.

Eddy, D.M.(1998). Performance measurement: problems and solutions *Health Affairs*, Vol. 17, Issue 4, pp. 7-25

Epstein, P.D. (1996).*Redeeming the promise of performance measurement issues and obstacles in the United States*, in Halachmi and Bouckaert, 1996 pp. 51-75.

Faris R.E.L & H.W. Dunham, (1965) *Mental disorders in Urban Areas, Chicago*, University of Chicago Press.

Forsythe, D.W. (ed.) (2001).*Quicker, Better, Cheaper? Managing Performance in American Government, Albany*, NY, November.

Gemeente Rotterdam (2008) ‘*JaarverslagBestuurlijkeHandhaving* 2007. Rotterdam.

Guthrie, J.E. & L.D Parker (1999).A Quarter of a Century of Performance Auditing in the Australian Federal Public Sector: A Malleable Masque, in: *Abacus,* vol. 35, nr. 3, p. 302-332.

Guthrie, J.P. & C. Schwoerer (1994).Individual and Contextual Influences on Self-Assessed Training Needs, in: *Journal of organizational Behavior*, vol. 15, nr. 5, p. 405-422.

Halachmi, A & G. Bouckeart (eds) (1996) *Organizational Performance and Measurement in the Public Sector*, Quorum Books, London.

Halachmi, A, (2011)Imagined promises versus real challenges to public performance management, in: *International Journal of Productivity and Performance Management*, 2011, vol. 60, nr. 1, p. 24-41.

Hood, C. (2007) Public Service Management by Numbers: Why Does it Vary? Where Has it Come From? What Are the Gaps and the Puzzles? *Public Money & Management*Vol.27 No 2, pp. 95 - 102

Hood, C. (1991) A public management for all seasons?*Public Administration*, Vol. 69 No 1, pp. 3 - 19

Hood, Ch., (1988) Articles - Individualized Contracts for Top Public Servants: Copying Business, Path-Dependent Political Re-Engineering -- Or Trobriand Cricket?, in: *Governance*, 1998, vol. 11, nr. 4, p. 443-462.

Judge, T.A. & Th. M.A. Welbourne (1994).A Confirmatory Investigation of the Dimensionality of the Pay Satisfaction Questionnaire, in: *Journal of Applied Psychology*; vol. 79, nr. 3, p. 461-466.

Kanter, R. Moss & D. Brinkerhoff (1981).Organizational Performance: Recent Developments in Measurement, in: *Annual Review of Sociology*, vol. 7, p. 321-49.

Kaplan, R. S. & D. P. Norton, (1992). The balanced scorecard - Measures that drive performance, in: Harvard Business Review, 1992, January-February, p.71-79.

Kluvers, R. (1998).The Development and Use of Performance Indicators in Victorian Local Government, in: *Accountability and Performance*, Vol. 4, No. 2, p. 61-77

Kouzmin, A. et al. (1999).Benchmarking and performance measurement in public sectors. Towards learning for agency effectiveness, in: *The International Journal of Public Sector Management,* Vol. 12, nr.2, p. 121-144.

Lohman, F.A.B., (1999).*The effectiveness of management information: a design approach to contribute to organizational control*, Delft.

Neely, A. et al. (1997). Designing performance measures: a structured approach, in: *International Journal of Operations and Production Management*, vol. 17, nr. 11, p.1131–1152.

Parker, L. & G. Gould (1999), Changing public sector accountability: critiquing new directions, in: *Accounting Forum*, Vol. 23 nr. 2, p. 109-135.

Pollitt, Ch. & G. Bouckaert, (2000) *Public Management Reform: A Comparative Analysis*, Oxford.

Spigelman, J. (2001). The 'New Public Management' and the Courts, in: *Australian Law Journal*,Vol. 75, p. 748.

Steers, R.M. (1975). Problems in the Measurement of Organizational Effectiveness, in: *Administrative Science Quarterly,* vol. 20, nr. 4, p. 546-558

Townley, B. et al. (2003).Performance Measures and the Rationalization of Organizations, *in:Organization Studies*,vol.24, p.1045-1071

Traub S. H. & C.B. Little, (1999) *Theories of Deviance,* Itasca, Illinois, F.E. Peacock Publishers.

Werner R.M &D.A. Asch, (2007).Clinical Concerns About Clinical Performance *Measurement, in: Annals of Family Medicine*, vol. 5, nr. 2, p. 159-163.

Young, C.W. & R. Welford, (1999).An Environmental Performance Measurement Framework for Business, In: Bennett; M, James, P; Kinkers, L (Ed.) *Sustainable Measures: Evaluation and Reporting of Environmental and Social Performance*, Sheffield, Greenleaf, p.98-116.