

Leistungsmessung: Je besser, desto missbrauchter

Information verändert die Politik – aber nicht immer zum Guten. Information ist nur dann fruchtbar, wenn die Bürger dank ihr die staatlichen Leistungen besser beurteilen können und die Politiker gleichzeitig starke Anreize haben, im Sinne der Bürger zu handeln. Das Problem von Benchmarking-Studien ist nicht, dass sie keine wichtigen Erkenntnisse liefern. Das Problem ist vielmehr, dass sie – gerade weil sie bei richtigem Einsatz wichtige Erkenntnisse liefern können – im politischen Prozess missbraucht und in partiischer und inflationärer Weise eingesetzt werden. Dadurch verlieren Benchmarking-Studien auch viel von ihrer Nützlichkeit.

Von Evaluationsstudien erhoffen sich viele nicht knackige Effizienzranklisten, sondern Empfehlungen zur Steigerung staatlicher Effizienz. Aufgrund der Anlage der Studien können jedoch oft nur schwer Handlungsempfehlungen abgeleitet werden, versuchen doch die Studien häufig, die Ineffizienz zu messen, aber kaum, sie zu erklären. Die Autoren der hier vorgestellten Studien argumentieren deshalb zu Recht, dass ihre Effizienzmasse und Rankings nicht als bare Münze genommen werden sollten, sondern dass schlechte Werte lediglich auf mögliche Probleme hindeuteten, die vertieft untersucht werden müssten. Ein Ergebnis der Studien ist aber dennoch herauszustreichen: Ausser in Versorgungsnetzen scheinen keine systematischen Grössenvorteile zu existieren; im Gegenteil: Gemäss der Studie von *Kersten Kellermann* steigen die Kosten pro Bürger mit steigender Kantonsgrösse. Die stereotype Forderung nach Kantonsfusionen aufgrund behaupteter Grössenvorteile ist damit ein weiteres Mal widerlegt.

Schwächen von Benchmarking

Die hier vorgestellten Studien leiden aufgrund ihres Forschungsansatzes insbesondere an folgenden, von den Autoren teils diskutierten Schwächen:

- Alle Studien erfassen nur die relative Effizienz. Sie können nur feststellen, ob eine Leistung teurer als vom günstigsten Anbieter, nicht aber, ob sie so gut wie tatsächlich möglich erbracht wird.
- Die meisten Studien fokussieren auf die Messung technischer Ineffizienzen; hinge-



Prof. Dr. Reiner Eichenberger

Leiter des Seminars für Finanzwissenschaft der Universität Fribourg, Forschungsdirektor des Center for Research in Economics, Management and the Arts (CREMA)

gen vernachlässigen sie allokativen Ineffizienzen. Mit teilweiser Ausnahme der Arbeit von *Kellermann* fragen sie nicht, ob die Leistung in Umfang und Zusammensetzung richtig erbracht wird.

- Die Studien fokussieren auf quantitative Aspekte der Leistungserbringung und müssen aufgrund fehlender Daten qualitative Aspekte vernachlässigen.
- Die Ergebnisse sind zum einen stark von der verwendeten Methode abhängig, was insbesondere die Studie zur Altpapiersammlung von *Jürg Kuster und Hansru-dolf Meier* eindrücklich illustriert. Zum anderen treiben das gewählte Sample und der Untersuchungszeitraum das Ergebnis, weil damit auch ganz direkt die Benchmark gesetzt wird.
- In der einen oder anderen Form zerlegen alle Studien die Kostenunterschiede in «legitime» Unterschiede aufgrund exogener Kostentreiber und «illegitime» Unterschiede aufgrund von Ineffizienz. Die gemessene (In-)Effizienz hängt somit davon ab, welche Faktoren als legitime Kostentreiber angesehen werden. Wenn etwa das Volkseinkommen als Faktor betrachtet wird, der gewisse kantonale Ausgabenunterschiede legitimiert, dann kann der Kanton Zug praktisch nur gut abschneiden. Denn dort sind die Einkommen so hoch, dass sie auch völlig unsinnige Ausgaben legitimieren. Ein anderes Beispiel ist die Studie zum Regionalverkehr von *Stephan Osterwald und Felix Walter*, in der die Anzahl Einsteiger pro Kilometer als legitimer Kostentreiber betrachtet wird, obwohl natürlich die Anzahl Einsteiger nicht exogen ist, sondern stark von der (In-)Effizienz der Leistung abhängt.

Diese Auflistung darf nicht als Kritik an den einzelnen Studien verstanden werden. Die Studien machen das Beste aus den verfügbaren Daten. Das Problem liegt vielmehr darin, dass Effizienzstudien gerade wegen ihrer Schwächen im politischen Prozess nur zu leicht missbraucht werden können.

Drohender Missbrauch

Viele Menschen denken, Benchmarking treibe die Leistungserbringer und die politi-

schen Entscheidungsträger zu mehr Effizienz. Schlechte Ergebnisse seien für sie unangenehm und hielten sie dazu an, akribisch nach Verbesserungen zu suchen. Das trifft aber nur unter günstigen Bedingungen zu. Die Gefahr ist gross, dass die Studien sogar eine gegenteilige Wirkung haben.

- Auch besonders gute, ausgewogene Studien haben oft nur wenig Einfluss, denn die Leistungserbringer reagieren völlig asymmetrisch auf ihre Ergebnisse. In den Kantonen und Gemeinden, die gut abschneiden, droht der Druck auf weitere Effizienzsteigerungen eher abzunehmen. Die Entscheidungsträger von schlecht beurteilten Leistungserbringern hingegen weisen die Ergebnisse der Studien zurück. Das fällt ihnen leicht, weil die Studien – wie oben gezeigt – sehr einfach zu kritisieren sind.
- Viele Studien werden unausgewogen und schlecht sein. Sobald Benchmarking im politischen Prozess eine gewisse Rolle spielt, wird es von den Evaluierten instrumentalisiert. Diese werden versuchen, ihre schlechten Leistungen mit verzerrten Benchmarkings zu beschönigen. Entsprechend obiger Kritik wird es ihnen leicht fallen, die Ergebnisse zu beeinflussen.

Die Ergebnisse solcher Auftragsstudien sind aus wissenschaftlicher Sicht wenig überzeugend. Selbstverständlich gibt es Methoden, um die Benchmarking-Studien entsprechend selbst zu benchmarken, etwa indem die Ergebnisse systematischen Robustheitstests unterzogen werden. Dabei kann sich aber nur zeigen, dass die Resultate tatsächlich wenig robust sind, wenn nur genügend viele Faktoren (Methode, Stichprobe, Untersuchungsperiode etc.) variiert werden. Das alles wird aber die Auftraggeber nicht daran hindern, sich mit parteiischen Benchmarking-Studien zu rechtfertigen. Dabei kommt ihnen entgegen, dass solche Studien nicht ganz billig sind. Auftraggeber sind deshalb zumeist staatliche Stellen, wohingegen regierungs- oder verwaltungskritische Gruppen die notwendigen Mittel nur selten aufbringen können.

In welchem Ausmass solche technisch orientierten Studien missbraucht werden können, zeigt sich etwa am Beispiel von Kosten-Nutzen-Analysen, die vorgeben, die Umwegrentabilität – d.h. die zusätzliche volkswirtschaftliche Wirkung – von Projekten und Ausgaben zu messen. Mit solchen Studien werden heute vielerlei staatliche Leistungen im Auftrag ihrer Leistungserbringer evaluiert. So hat das Sozialdepartement der Stadt Zürich zeigen lassen, dass Kinderkrippen und die Sozialhilfe volkswirtschaftlich nützlich sind.

Andere Ämter und Interessengruppen haben die volkswirtschaftliche Nützlichkeit des öffentlichen Verkehrs, des privaten Verkehrs, von Flughäfen, der Euro 2008, des KKL Luzern, von unzähligen Museen und Kulturinstitutionen usw. zeigen lassen – zumeist in der Absicht, mehr öffentliche Mittel zu erhalten. Natürlich müssen für die erwünschten Ergebnisse oft hanebüchene Annahmen getroffen werden. Das aber stört die Auftraggeber wenig. Denn mit solchen Studien ist es ähnlich wie mit Anwälten und Bodyguards: Sie nützen einem, ob man Recht hat oder nicht.

Der richtige Ansatz: Bessere Anreize für Politiker

Die bisherigen Überlegungen lassen vermuten, dass der zunehmende Einsatz von Benchmarking-Studien den Bürgern nur wenig hilft, die staatliche Leistungserbringung treffsicherer zu beurteilen. Demzufolge werden sie die politischen Entscheidungsträger auch nicht dazu bringen, dem Willen der Wähler entsprechend zu handeln. Zur Effizienzsteigerung im öffentlichen Bereich gibt es deshalb nur ein Mittel, nämlich die Verstärkung der Anreize für politische Handlungsträger, im Sinne der Bürger zu entscheiden. Dafür brauchen wir nicht nur eine wirksamere direkte Demokratie (z.B. das Finanzreferendum auf Bundesebene) und eine Stärkung der föderalistischen Institutionen (etwa mehr Gemeindeautonomie mit grösserer finanzieller Eigenverantwortung), sondern insbesondere auch die Schaffung von unabhängigen, volksgewählten Kontrollgremien, die – ähnlich wie kommunale Rechnungsprüfungskommissionen – die Politik auf Kantons- und Bundesebene begleiten. Solche Rechnungsprüfungskommissionen, die über nichts selbst entscheiden, aber alles kritisieren und evaluieren sowie Gegenvorschläge bei Volks- und Parlamentsabstimmungen ausarbeiten können, hätten denn auch weit bessere Anreize als die heutigen politischen Entscheidungsträger und die Verwaltung, Benchmarking in unverzerrter, unparteiischer und wirklich informativer Weise einzusetzen. ■