Benchmarking carrots and sticks: developing a model for the evaluation of work-based employment programs

Castonguay, J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
5. Summary and conclusions on social benchmarking

The previous chapter has shown that a very rich amount of information will be compared within this benchmark model, which brings the discussion back to the use of benchmarking for evaluating social policy. As already explained in chapter two, benchmarking was first introduced by the private sector and was used to compare the production chain within an industry. Five different types of benchmarks were identified, where each focused on one section of this production chain, such as production costs or production processes. In the public domain, with many countries looking to improve their active labour market policies, several international comparisons have been undertaken in this field. Learning from best-practices has become more and more important in policy-making, and international evidence provides a mean to evaluate how rather different systems compare to each other in terms of performance. However, such international comparisons of labour market programs are not straightforward since many elements of differences and similarities need to be taken into account. Early adapters of benchmarking in the employment policy domain based their model on benchmarks which where developed for the private sector, and thus mostly focusing on one single aspect of the policy-chain. This resulted in simple benchmarks, which often failed to show how the ranks of each public program were being attained. Many thus saw benchmarking as a trend through which policy-makers attempted to increase the visibility of their efforts in achieving effectiveness and efficiency, without using that benchmark in order to improve their program. Benchmarking is thus often seen as window-dressing, where giving the impression of working on the effectiveness and efficiency of the programs was what was being sought, more than improving the programs themselves.

The research question of this benchmark asks how benchmarking can help in evaluating current active labour market policies. What is clear is that simple benchmark models in which only a few impact indicators are being compared will most likely not contribute to a better understanding of employment policies. Nevertheless, by building a comprehensive social benchmark model based on the entire policy-chain, this benchmark will be able to point out which elements of the input, the process, the output and the external factors where crucial in reaching the impacts in each program. Active labour market policies will then be able to be improved, based on the determinants of success and failure which where discovered through the benchmark.
It should thus be clear that social benchmarking is not a method which is an easy "quick fix", since a deep comprehension of a broad array of elements concerning not only the programs themselves but also the context in which the programs takes place is in fact necessary in order to fill-in all those indicators. Data sources will be diverse, from legislations on social security and employment protections and official policy statements from the government, to qualitative and quantitative evaluations performed by outsiders to the program.

The strength of benchmarking clearly lies in this ability to use both qualitative analysis as well as quantitative analysis in order to make conclusions on best-practices. While social benchmarking could be seen as a purely quantitative method, this is not the case for the benchmark being developed here. Indeed, if a benchmark contains a large number of cases as well a small number of indicators, quantitative analysis can then be used to test the links of causality between the different indicators. This is the method which was used with great success by Mosley and Muller (2007) in their benchmark of 141 German Public Employment Service offices across the country. However, this benchmark favours an analysis in which fewer cases are analysed in detail with respect to many elements of their policy chain. This way, a much deeper understanding of the many facets of the programs can be achieved, which surely will benefit the analysis of the results of the programs as the benchmark is aggregated to its final components.

Clearly, the social benchmarking model developed in this research finds itself at the intersection of two different research methodologies. As already said, it is much more detailed in its accounts of influential factors than what is mostly done in quantitative analysis. This way the model will be able to highlight best-practices while being aware of a large amount of elements which work together to create a successful program. On the other side, the benchmark model goes much further in quantifying and aggregating information than most qualitative research do. This way, the model will be able to sum-up the large amount of information gathered in a concise manner. More than a technique to sum up qualitative information, social benchmarking then allows to draw some conclusion on overall level of performance and show how different inputs and processes may lead to different levels of efficiency and effectiveness. Causality is then assessed in a much stronger way than in traditional qualitative research.

Hence, while the 25 indicators of this benchmark model might seem too much for those used to quantitative analysis based on a few causal relations, these might seem like a shallow depiction of the reality of
employment programs for those used to qualitative analysis. However, benchmarking does offer the promising possibility to act as a bridge between quantitative and qualitative analysis, combining the strengths of numbers with that of a broader view on what creates successful programs. Furthermore, social benchmarks do provide a platform through which both types of research can meet. On the one hand, strong quantitative research is needed in order to make assumptions on which the rankings of the indicators can be based. But on the other hand, a broader view of employment programs is necessary in order to make a proper choice on which indicator could be useful within each category of the policy chain. This choice should clearly be inspired by qualitative research on the realities of employment programs, both for individual participants as well as for policy-makers and managers.

The next part of the research has two objectives. On one side, the intention of the benchmark exercise will be to test the benchmark model which was developed in these last chapters. In other words, “the proof of the pudding is in the eating”. Claims with respect to the usefulness of benchmarking for evaluating labour market policies will be able to be assessed not only theoretically but also in an applied context. On the other side, this research has also for objectives to make recommendations on how to improve work-based employment programs and assist policy-makers in making the right choice when designing their program. Hence, whether this social benchmark model works and whether work-based employment programs also work will be investigated next.