

File ID 146376
Filename 11: Benchmark synthesis

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type Dissertation
Title Benchmarking carrots and sticks : developing a model for the evaluation of
work-based employment programs
Author J. Castonguay
Faculty Faculty of Law
Year 2009
Pages 388
ISBN 978 90 5629 575 2

FULL BIBLIOGRAPHIC DETAILS:

<http://dare.uva.nl/record/315571>

Copyright

*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.*

11. Benchmark synthesis

A number of conclusions have already been drawn from the analysis presented so far. First, there is a wide variation in the type of input and process involved in work-based employment programs in the countries included in this research. This means that one should be careful when speaking of “workfare” or work-based employment programs using generalisations. In addition, the variation found in the input-benchmark was not the same as found within the process-benchmark. Actually, both differed greatly from each other in terms of which program could be expected to perform best when looking at output and impacts. Second, it was made evident that there is also a wide variation in the results of those programs, both in terms of their outputs and their impacts. Answering the question of whether work-based employment programs do succeed in assisting the unemployed in finding jobs thus really depends of which programs is considered. Third, while external factors are important in understanding the background and results of social programs, the amount of variation found in the indicators of this benchmark was smaller than expected. Nevertheless, some outliers with respect to either of the economic, legal or political context will possibly provide for some explications on diverging design and results.

Conclusions on the relationships between all these separate indicators remain to be drawn in this chapter. By looking at the links between inputs, process, outputs, impacts and external factors, the efficiency and effectiveness of each program will be determined. Moreover, the analysis of the inputs and process with respect to the outputs and the impacts will allow specifying the determinants of success or failure of the programs in this benchmark. Through this, recommendations can be given in the next chapter on how to best design a work-based employment program in order to achieve the aim of fostering a return to the labour market.

11.1. Effectiveness

Measuring the effectiveness of a program means measuring to what extent it has reached an objective. This objective can be either determined by the evaluator, or be taken directly from the program to be evaluated. In this benchmark, a combination of both will be done. First, chapter 3 discussed how the intervention strategy of work-based employment program is meant to increase the return to work of the unemployed. A high return-to-work ratio for the participants in the program is thus chosen in this benchmark as the main objective for the measurement of effectiveness. Nevertheless, it was also discussed how this intervention

strategy is made of three different dimensions: the willingness to work, the ability to work, and access to work. These objectives relate more to the output of the program and can also be part of the objective of the programs in the benchmark. An evaluation of the extent to which each program reaches the objectives it made explicit in policy documentation will thus also be done.

The effectiveness of the programs with respect to the return to the labour market has thus already been discussed when presenting this indicator in chapter 9. The most effective program was shown to be the Employment Option within the New Deal for Young People, with more than 55% of its participants in a job by the end of the program. The Dutch Work First programs came in second, with on average 45% of the participants joining the labour market. The Voluntary Sector and the Environmental Task Force options of the New Deal for Young People and the Work for the Dole programs then came in third, with somewhere between 25% and 30% of their participants finding a job. The ND25+ and the Temporary jobs program come next with a return to work ratio of one fifth. The least effective programs with respect to return to work are the RMCAS and Ontario Works, with only 1 participant in 10 leaving the work-activities for the labour market.

Benchmark 1 had showed that return to work was a strong objective in the Netherlands, Canada and the United Kingdom. Within this perspective, the good results of the Work First in the Netherlands can be said to be showing a high level of effectiveness, not only according to the objective of this evaluation but also according to the internal objectives of the program. On the other side, the very poor results of the Ontario Works program does highlight the great ineffectiveness of this program, even taking its own objectives into account. Internal effectiveness with respect to the New Deal programs diverges depending on which program is being looked at. This is due to the fact that all New Deal programs share the same general objective while results vary greatly between its different components. The New Deal thus clearly reaches its objectives with the Employment Option of the NDYP, and much less with the Work Experience / Placement IAP of the ND25+. Nevertheless, with respect to these last programs as well as for the VS/ETF Options, the objectives of these specific components of the New Deal programs did bring some nuance in the focus on return to work. In fact, all these options could also be shown to intent to provide work experience to the participants, and especially in the ND25+, allow the participants to get back in touch with the realities of working (see Hasluck and Green 2007, p. 45). Hence, the indicator with respect to the objective of the program could be nuanced here away from a strong focus towards return to work, making internal

effectiveness less low than it would first appear. Nevertheless, external effectiveness according to the objectives given in this benchmark remains at its lower level. It should also be noted how it was shown that within the ND25+ the Subsidized Employment program, although too small to include in the benchmark, was much more effective in assisting participants in being employed by the end of the program. Its outflow to work ratio was closest to that of the Dutch Work First programs.

It is true that the return to work objective is less central to the Work for the Dole program. Nevertheless, even the Department for Employment and Workplace Relations evaluates the Work for the Dole program by looking at how many persons are employed after completion (see DEWR Annual Reports as well as DEWR, 2006, 2002, and Burgess et. al. 2000). In order to provide a complete evaluation of both the internal and external effectiveness, the Work for the Dole will also be evaluated against its own objectives next. Concerning the two Swiss programs, both the RMCAS and the Temporary Jobs programs can also be evaluated on its effect on the return to the labour market of its participants, since this is also considered as an objective of the program by official governmental evaluations (see Cunha et. al. 2002, Fluckliker and Vasiliev, 2003 and CEPP, 2002). It is nevertheless the case that the Temporary Job does not have the objective to assist the return to work of its participants. On the contrary, its objective is allowing unemployment claimants to be eligible for a second term on the benefit. Since about 80% of its participants do end-up claiming federal UI after having taking part in the program, it can be said that, according to its own objective, the program is rather effective. However, seeing the fact that its design fits the design of programs which do have the objective to help claimants returning to the labour market, it is also possible – and interesting – to evaluate the program against this objective. One could possibly find out that a high rate of return to the labour market is achieved through this design, while being unintentional. Unfortunately, this turned out not to be the case, and as already discussed earlier, this program has one of the lowest rate of return to the labour market. An explanation for this will be discussed later on.

Effectiveness with regards to the intervention-strategy of each country has to be related to the sub-objectives of benchmark 1. This benchmark shows that increasing the skills of the unemployed is an important objective of the New Deal programs as well as the Dutch Work First projects, although the latter to a lesser extent. The Work for the Dole program also aimed at somehow increasing skills, although this was more indirectly intended, through providing work experience and not formal skills and qualifications. The objectives of increasing skills levels is

evaluated against the output indicator with respect to the amount and quality of the training provided through the work-based employment programs. Benchmark 15 shows that the training elements of the New Deal program are in fact the best ones in this benchmark, as they both scored a 5. The objective to increase the skills of the unemployed has thus been achieved by providing all New Deal participants with a substantial amount of formal training in the work-based employment programs. The New Deal can therefore be said to be very effective according to both the general objective of increasing exit-to-work and its more specific intervention-strategy aimed at increasing skills levels. The Work First projects in the Netherlands had received a score of 3 with respect to training, and in the view that the objective of increasing skills was not as strong as in the UK, it does not strike out as being particularly ineffective with respect to this objective. The Work for the Dole had received the lowest score with respect to formal training, but as explained earlier, the improvement of skills was not meant to be done through formal training but rather through work experience. Hence, it cannot be said to be ineffective with regards to its own training objective.

The aim of creating a mutual obligation was central to the Work for the Dole program, the Ontario Works programs and the RMCAS programs. In these three programs, the requirement to participate in work-activities was justified by the need for the benefit claimants to give something in return for the financial assistance they receive. Naturally this objective is being achieved *de facto* in these programs, by designing work-based employment programs as part of the benefit scheme. Nevertheless, effectiveness here can be evaluated by looking at the coverage of the work-based employment system. The reasoning behind this is that a program with a very low coverage would thus not succeed in requiring this mutual obligation from most of its claimants. As can be seen from the radar charts in the output-benchmark, the highest coverage was found in the Ontario Works programs as well as the Temporary Jobs program. In both these program, about half the benefit claimants participated in the work-activities. Thus according to the intervention strategy of the Ontario Works program, it does succeed in being effective in achieving its aim to create a high level of mutual obligations. This contrasts with its very low effectiveness with respect to its objective to increase the outflow to work of its participants. Coverage was at a medium level in both the Work for the Dole program and the RMCAS, making these programs less effective than Ontario Works with respect to creating mutual obligations. Also interesting is the fact that coverage is very low in those programs that did not focus on mutual obligations, the New Deal and the Netherlands.

The objective to increase access to work for benefit claimants was in none of the program a central element of its intervention strategy. Nevertheless, the official objectives of the Work for the Dole do state that the program aims to help the unemployed build networks, which can then be of great importance in finding a job. The extent to which this is being achieved was discussed in the benchmark on supervision and guidance, and in specific the supervision provided by the Community Work Coordinators. It was shown there that there is substantial evidence from surveys that the high commitment level of the Community Work Coordinators (CWCs) is the key element in the program's results (see Nevile, 2003). Many CWC's were reporting that they attempted to find jobs for their participants through potential employers they knew or through their own organisation. However, such assistance in accessing the labour market cannot be said to be part of the official design of the programs as it is neither being enforced by the Department nor by Centerlink. This lack in the enforcement of this objective is even weaker considering the fact that there are no financial incentives in the contracts between Centerlink and the CWC's with respect to participants finding jobs. The overall score for the output-indicator was moreover much diminished by the very low level of supervision and guidance from Centerlink itself. The conclusion which can be drawn from this is that this objective is being attained in the program, but not resulting from transferring this objective to the implementing bodies of the program, but from the goodwill of those implementing it. Such an effectiveness level is thus very fragile, as it is neither monitored nor enforced by the government body responsible for the program.

One last objective should also be discussed with respect to the Work for the Dole program, being the objective to increase the self-esteem of the benefit claimants. Carson, Winefield, Walters and Kerr (2003) reviewed two major studies on the experience of Work for the Dole participants performed by the government department responsible for the program. These studies surveyed participants on the topic of self-esteem, psychological well-being and work attitudes, and presented noteworthy results on the extent to which the objective with respect to building self-esteem is being achieved. The program was found to have no effect on the level of self-esteem of the participants, unless the program did lead to a sustainable job deemed "worthwhile" by the participants. Carson, Winefield, Walters and Kerr also note how this increase in self-esteem is thus not directly created by the program, but appears only after completing the 6 months program and exiting for employment. In addition, these studies showed that attitudes and commitment with respect to work were actually high before the claimants started the program, and that these remained high throughout the program, which

according to the authors did contradict the assumptions about the lack of motivation of the benefit claimants. On the other side, the program was found by Carson, Winefield, Walters and Kerr to have a positive effect on the psychological well-being of the participants, mostly through preventing boredom and social exclusion. All in all, with respect to its own objective of raising self-esteem for participants, the program can be concluded to be ineffective. Since only one third did find a job through the program, and possibly a smaller proportion did find a job which was worthwhile to them, the large majority of the participants did not see the program have any influence on their self-esteem.

To sum up the findings so far on the effectiveness of the programs in this benchmark, efficiency with respect to return to work was found highest in the New Deal for Young People Employment Option and the Work First programs. Both these programs also had a very strong focus for this objective in their program, so these can be said to be efficient with respect to this benchmark's criteria for efficiency as well as with respect to its own objective. Efficiency was found to be low in the RMCAS program, Ontario Works, the Temporary Jobs, and the New Deal 25+. Return to work was a strong objective of the Ontario Works programs and to a lesser extent in the New Deals programs, but not for the other two. The Work for the Dole program and the VS/ETF Options of the NDYP were found to also be performing rather low, but still better than the those four programs just mentioned above. With respect to their own objectives, the objective of raising skills was found to have been reached in the UK. The Ontario Works program did reach the objective of creating mutual obligations, and this was less the case for the Work for the Dole program and the RMCAS, who also shared this objective. The Work for the Dole program also failed to raise self-esteem for its participants, and the effectiveness achieved with respect to building networks was a weak one. The temporary job program had a very different objective, that of allowing the participants to become eligible for a second period of unemployment insurance, and because of its high coverage it can be said to have achieved this objective.

External Factors

Effectiveness should however be looked at in the light of the external factors, as these could explain why some countries are performing better than others in this benchmark. It was already explained that impacts should optimally be corrected for differences in external factors as to truly portray the differences they make in the increasing outflow to work. Unfortunately, data on net impacts were not available for most of the programs in this benchmark. The second-best solution to this is to look at

the return to work ratios in comparison with the indicators of the external factor benchmark. Although this will not allow a quantitative correction of the data measured under the impact benchmark, it will make it possible to at least qualitatively acknowledge those factors.

The problem with most of those external factors is that their effect on the rate of outflow to work is not straight forward. This has already been discussed in both chapter 4 and in its benchmark chapter, so this discussion will not be repeated here. In general, it can be acknowledge the higher unemployment rate in Canada could contribute to making it more difficult for Ontario Works participant to find jobs. However, this rate has been decreasing from the beginning of the 2000's onward (Statistics Canada, 2007). That is not to say that this factor alone can be made responsible for lower effectiveness, and that all things being equal in the economic context, this program would perform just as well as the NDYP Employment Option and the Dutch Work First programs. In fact, it can be shown that programs within the UK all had rather diverging degree of success, pointing to other factors than external factors as a main explanatory factor for good results. The next section will show that input and process also provide some explanations as to why this program performs worse than the others.

Other impacts

Effectiveness has so far been relative to achieving high rates of outflow to work or a high level of output which eventually leads to high rates of outflow to work. As discussed throughout chapter 9, very little data is available which would allow the comparison of other important impacts of work-based employment programs. In particular, different approaches to work-based employment programs could have varying impacts with respect to job sustainability and on the preventive effect they have on entry into the programs. But without this data, this benchmark is unable to make conclusions on effectiveness with respect to these objectives, whether or not these are aimed at by the programs. In addition, it was also mentioned how total caseload reduction, without specifying that this needs to be through movements into employment, can also be an important objectives of employment programs. Indeed, this was shown to be the most central within the Dutch Work First programs. Chapter 9 had concluded on this that this was indeed the case that that total caseload reduction was highest in that country, such that with respect to over caseload reduction, the Netherlands can also be said to have reached its objective.

It should thus be clear that “success” and “performance” will mostly be based here on one single impact indicator, the outflow to regular work. Efficiency and effectiveness should therefore be understood as being with regard to what can be said to be the central aim of work-based employment program. In other words, this benchmark cannot provide a measure of the *overall* success of work-based employment program. Nor does it attempt to measure all impacts these programs have on the individuals, since most are unmeasured. That is not to say that these program might not have a significant impact on other aspects of the lives of participants, such as alleviating social exclusion, improving psychological well-being, providing a stepping-stone for reaching out to other social provisions such as social housing, or improving other areas of life such as financial stability, interpersonal relationships, and much more. This area of research which is taking a broader perspective on the problem of unemployment is a very promising one, and as the results of this research agenda become available, the social benchmark model developed here will be able to include them in its evaluation criteria.

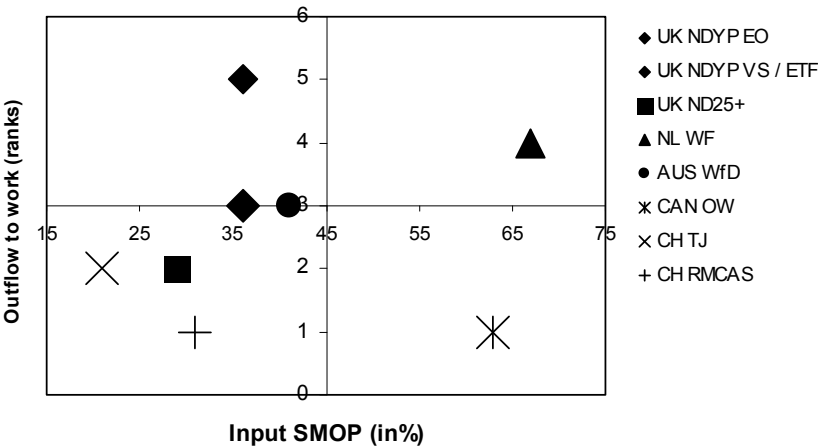
11.2. Efficiency

Efficiency is defined as the extent to which the results of the program match up its level of investment. The definition of “investment” has to be interpreted differently than usual in this benchmark, since this will be measured by the input which relates more to the initial resources and parameters of the program than financial costs. The question is then whether countries which had a low effectiveness level actually did not present the initial conditions within the program which could foster high return to work ratios. On the other side, analysing whether programs with high level of input also achieved the best results allows to test the hypothesis made when ranking each countries. Inputs were ranked based on their expected effect on results, with a high rank expected to lead to better results. Incongruence between the level of inputs and impacts could thus be explained two factors. The first is that other parts of the policy chain would have a stronger influence than the inputs. And thus the second one being that rankings did not follow the appropriate hypothesis as to the extent to which a certain input indicator is expected to lead to better result. Such a possibility was in fact discussed with respect to two of the seven input indicators, namely the level of the benefit and the harshness of sanctions. Special attention should thus be given to these two indicators when looking at the efficiency of the programs.

From the radar charts and the surface measure of performance of the input-benchmark, it is apparent that the Dutch Work First programs and the Ontario Works programs have the highest total level of input. In other words, the initial conditions in place are most likely to make sure the program will succeed. The lowest level of input is found in the Temporary Jobs programs. The other four programs have total levels of input close to each other, at about half to two-thirds of what is found on the aggregate level in the two best programs.

By crossing over the two indicators of impacts and input, a four quadrant typology is then formed, as shown in figure 11.1. The upper portion of the figure is where effective programs can be found. The y-axis shows the overall performance measure for the inputs, such that the right side of the figure contains the programs with the highest level of input.

Figure 11.1 Input-efficiency in four quadrants



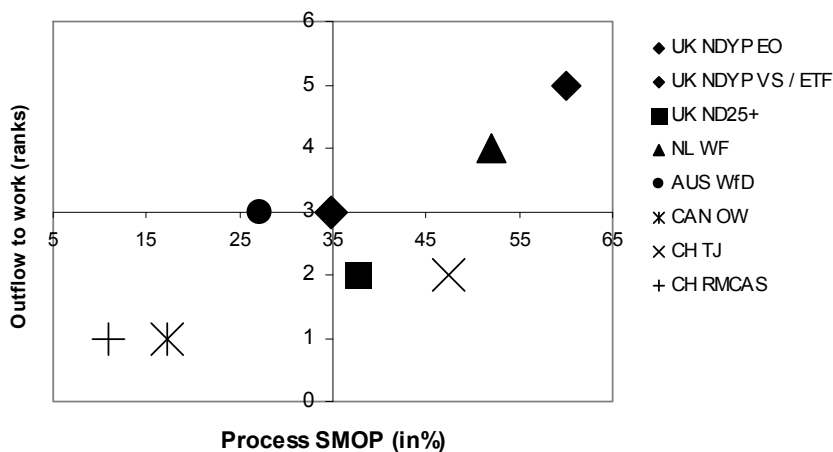
In the first quadrants on the right and upper side in the figure both the inputs and the impacts are high, meaning that the program is being efficient in transforming input into impacts. In other words, investing in a high level of initial conditions for the program has paid out and the program has been successful. This is thus the case for the Dutch Work First projects, which are the only programs in this quadrant. The situation is very inefficient in the Ontario Works program, since it also shows a high level of input, but this program is found in the lower side of the figure where impacts are low. This program is therefore the most inefficient program in this benchmark. The situation in both Swiss as well

as the ND25+ programs is different, since the low impacts were reached through low level of initial conditions which were needed for a successful program. Such a situation cannot be said to be inefficient, since their low impact is compensated by low input, such that resources and efforts cannot be said to have been wasted on a programs with bad results. On the other side, the NDYP Employment Option program has reached the best results with a similar level of inputs. This makes this program the most effective of the benchmark, because it succeeded in reaching high rates of return to work even though the initial conditions to do so were not optimal. At last, the Work for the Dole program and the NDYP Voluntary Sector and Environmental Task Force Options are found very close to the centre-left of the figure, meaning that (slightly lower than) average inputs created average impacts.

The ranking of the input-indicators was done based on the hypothesis that a high rank would result to high impacts, so it is not surprising to see a more or less linear relationship between these two elements of the policy chain. However, the very high level of efficiency measured in the NDYP Employment Option might also point towards some rankings not actually following the right direction. Hence, instead of having low levels of input, the UK would actually be found to have high levels of input. One way to solve this is through looking at the radar charts for the inputs. What can be seen is that while the size of the overall input level is the same between these programs, their shape is different. This indicates that explanation for the links between impacts and inputs must be looked at within single indicators. Looking at those single determinants of success will be done in section 11.3 next.

Efficiency can also be defined in terms of the process. Process-efficiency measures the extent to which the choices in the process resulted in a high level of impacts. Here again, a four quadrant graph can be used to show the different combinations of process level and impacts in the programs of this benchmark. High outflow to work can still be found in the upper part of the graph, and low outflow to work in the lower part. Process is then plotted on the left-right axis, with high overall scores in the process benchmark on the right, and low overall score on the left. Here again, the top-right quadrant can be said to show a high level of process-efficiency, since the high level of process was matched by a high level of impact. Consequently, the NDYP Employment Option and the Dutch Work First programs can be said to be efficient, and the Temporary Job and the ND25+ programs to be inefficient.

Figure 11.2 Process-efficiency in four quadrants



For the Ontario Works program and the RMCAS, the overall scores on the process are low and low impacts are thus realised. This situation here also cannot be said to be inefficient, since they can be expected from the choices being made in this program. The Work for the Dole program and the NDYP VS and ETF Options on the other hand do achieve better results than the Ontario Works program and the RMCAS, through slightly better process levels than these two programs. In fact, it is found somewhere in between the NDYP Employment Option and these two programs, meaning that it achieved average results with average process scores. The Work for the Dole and the NDYP VS/ETF Options were also found to be in this middle position with respect to input-efficiency. In general, it can also be seen from the two previous figures that, as compared to the input-benchmark, the relationship between the process and the impacts is much more linear.

At last, output-efficiency can also be measured through this benchmark. The intervention-strategy explained how increasing the ability, willingness and access to work of the participants will increase the chances they find a job on the labour market. Comparing the amount of training and supervision/guidance to rates of outflow to work will thus allow measuring the output-efficiency for each program. When taking the average ranks of the training indicator and the supervision and guidance indicator, the highest ranks are found in the two New Deal programs, with the NDYP receiving a 5 and the ND25+ receiving a 4.5. The Dutch Work First projects receive a 3, and all the other programs receive an average score of 2. This makes the Employment option and the Work

First programs efficient because of high output and high impact. The other New Deal programs and then inefficient because they have high output but low impact. The Work for the Dole can be said to be very efficient side since it score also a 2 on the output, but had better results than the programs with a 2 on the output indicators. The two Swiss programs and the Ontario works program are efficient through both low output and impact.

Coverage was the third elements in the output benchmark and should also be put in relation to the impact of the programs. Indeed, a low coverage would point to the fact that a program does not reach many unemployed persons in the benefit, and that the efficiency and effectiveness which is being achieved only benefits a few individuals. The work-based employment program, in some case as part of a larger program, thus will not contribute a lot to decreasing caseloads in general. Interestingly, the programs which have the best scores with respect to efficiency and effectiveness also have the lowest coverage rates. As a result, even though a large part of the participants in the NDYP work-based options, the ND25+ work-based intense activity period, and the Dutch Work First programs do actually find a job through these programs, they are not likely to have a great impact on decreasing the total unemployment caseloads. On the other side, programs such as Ontario Works and the Temporary Jobs programs have the potential to have a larger impact on caseloads, since many claimants participate in their work-activities. But these two programs are also having a low rate of outflow to work, meaning that their total effect on caseloads still is not very large.

Input-efficiency, process-efficiency and output-efficiency thus all contribute to the final results each program is able to reach. Table 11.1 on the next page summarises for each program their level of efficiency on each of these elements of the policy-chain.

The efficiency of the New Deal for Young People Employment Option and the Dutch Work First programs can be mostly explained by their high efficiency level all throughout the policy-chain. Efficiency is also reached by the RMCAS program, although this is because low impacts where achieved through low levels of input, process and output. This program can however hardly been seen as successful, as shown by its low effectiveness level. The Work for the Dole does not strike out as being particularly successful, but since this program also only reach average levels of inputs, processes and output, it can also be said to be efficient.

Table 11.1 *The efficiency of the policy-chain*

	Input-efficiency	Process-efficiency	Output-efficiency
United Kingdom			
<i>NDYP EO</i>	Very efficient	Efficient (H/H)	Efficient (H/H)
<i>NDYP VS/ETF</i>	Average	Average	Inefficient
<i>ND25+ IAP</i>	Efficient (L/L)	Inefficient	Inefficient
The Netherlands			
<i>Work First</i>	Efficient (H/H)	Efficient (H/H)	Efficient (H/H)
Australia			
<i>Work for the Dole</i>	Average	Average	Very efficient
Canada (Ontario)			
<i>Ontario Works</i>	Inefficient	Efficient (L/L)	Efficient (L/L)
Switzerland (Geneva)			
<i>Temporary Job</i>	Efficient (L/L)	Inefficient	Efficient (L/L)
<i>RMCAS</i>	Efficient (L/L)	Efficient (L/L)	Efficient (L/L)

Note: H/H = high (input / process / output); and high results

L / L = low (input / process / output); and low results

The same can be said about the NDYP Voluntary Sector and Environmental Task Force Options, although these were found to be inefficient with respect to output. The Temporary Job program has one of the best overall ranking for its process, which would have lead to believe that this program could have had one of the highest impacts. However, the combination of low input and low output seem to explain much more the low level of impact of this program.

The New Deal 25+ was in the same situation, with low input and process also relatively high, although much closer than the average. In the Canadian program, the opposite is true, where a high level of input was not matched by high ranks on the process indicators, resulting in poor impact level. Findings on these two programs point to the need to match input to output in order to reach optimal levels of impact. In other words, high inputs or high process alone are not sufficient in order to guarantee the success of the program.

11.3. Determinants of success: the carrots-and-sticks index

Not only total levels of input, process and output are important in understanding the source of performance gaps in the programs, but single indicators can also provide interesting insights in the links between

design and results. As was noted in each second section of the previous chapters, the radar charts were used not only to illustrate the total size of each element of the policy chain, but also to visualise the differences in focus in the input, process and output. Although these differences were mentioned in each of those sections presenting the radar charts, these were not linked directly with the results of the programs, or the other elements of the policy chain.

Target Groups

The net-impact of labour market programs is usually also calculated by correcting for the characteristics of the participants in the programs, which might not be representative of the total population. Indeed, it could be so that the selection of the participants for the programs (either through mandatory or voluntary allocation) is biased towards persons with more or less barriers to work. A program with a target group which is very close to the labour market would therefore have better results than other programs, even though other elements of the policy chain would be similar. This was indeed the hypothesis formed when ranking the indicator of the target group in the input benchmark (Benchmark 2).

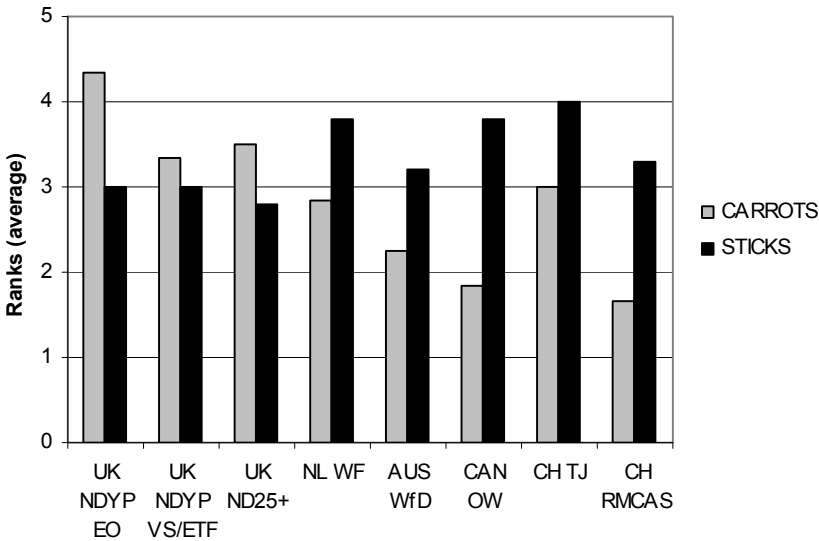
Looking at only the rankings on the target groups and the outflow to work, the hypothesis that young and short-term unemployed persons will lead to better results is only partially proven. In line with the hypothesis, target groups where favourable to good results in the NDYP Employment Option and the Dutch Work First programs. On the other side, the ND25+ and the two Swiss programs had the most difficult target groups and results were also lower in line with the hypothesis. However, the best score for target group was found in the Work for the Dole program, who only reached average results in terms of outflow to work. Similarly, the target group of Ontario Works was ranked in the middle of the range, but this program had the lowest score in the benchmark (together with the RMCAS). Thus, despite having a relatively easier target group in terms of age and unemployment record, this program was not able to reach good results. Looking back at the discussion on the Ontario Works target group, it was mentioned how lone mothers make-up a third of the participants. Perhaps the presence of a large group of lone mothers could mean that the target group of the Ontario Works program was overrated, and should have been scored lower. Lone mothers were either completely excluded from the other programs (in the UK and Australia) or not a specific large part of the programs (in the Netherlands and in Switzerland).

If one would want to correct the indicator of the outflow to work for the target groups in the programs, this would mean that the ND25+, the Temporary Job program and the RMCAS should receive slightly higher ranks than their scores of either 1 or 2. Doing so would bring these programs much closer to the NDYP VS/ETF Option and the Work for the Dole program. When looking back at the input and process efficiency, this would actually make the correlation between input and impact and between process and impact much more linear. Especially in the case of the ND25+ and the NDYP VS and EFT Options, the target group explains very well how a very similar process and input did result in different outflow to work. Nevertheless, as shown from the Ontario Works program and especially the Work for the Dole program, only focusing on target groups will not guarantee that high rates of return to work will be reached.

The carrots-and-stick index

The indicators of the input and the process were ranked following general theories on their micro and macro influences on employment. Micro influences were also explained in chapter 3 to be divided into two categories, positive incentives and negative incentives. The carrots and sticks metaphor illustrates well how the positive incentives are meant to reward appropriate behaviour and the negative incentives are used to punish undesirable behaviour. These two categories of mechanisms can also be seen as on one side the pull-factors and on the other side the push-factors, which “activate” the unemployed in the “direction” of employment. Making this distinction might thus shed some light on the performance of different programs in the benchmark, on a different level than that of the policy-chain. Three input indicators were already shown in chapter 4 to be stick indicators. These are the sanctions level and duration, the activation conditions, and the benefit level. To this category of negative incentives can be added two process-indicators, namely an early timing of the start of the program and a high number of hours in the work-activities. High rewards, short programs, private sector working-environments, fair sanctioning procedures, high level of training and high level of supervision and guidance can then be seen as positive incentives within the programs. Figure 11.3 shows these two categories of indicators, for which an average was calculated in order to give an overall measure.

Figure 11.3 Positive and negative incentives in each programs



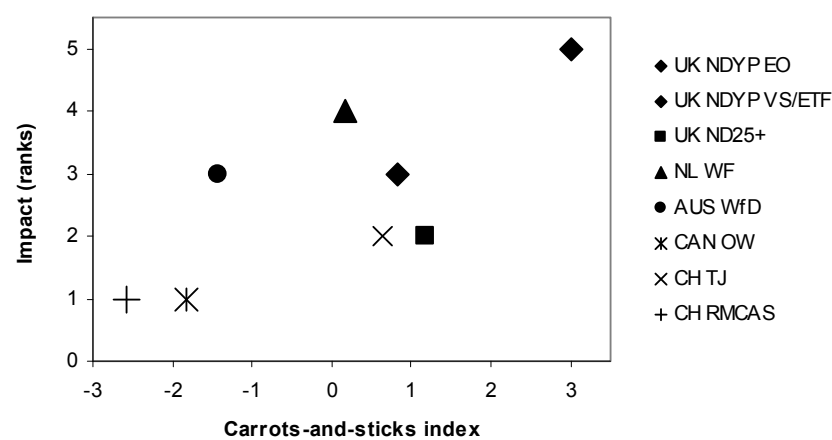
Looking back at the radar chart on inputs, one can see that the explanation for the lower overall size of the inputs in the UK was mostly due to low score on two out of the three stick-elements in that part of the policy-chain (and with respect to ND25+ the target group as discussed above). In fact, figure 11.3 shows that the average rank for the stick-indicators in the New Deal programs is in fact the lowest in the benchmark. This figure also shows that there is an important variation in the average ranks for the positive incentive in the different programs. These are by far the highest in the NDYP Employment Option, and lowest in the two less effective programs, namely Ontario Works and the RMCAS. What is also apparent from this figure is that within the average ranks for the positive incentives are much higher than then average ranks for the negative incentives in the New Deal programs. This is not the case for the other programs in the benchmark.

In order to further explore the relationship between positive incentives, negative incentives and the outflow to work ratios, an index can be constructed summarising the information in table 11.3. The first component of the index is the total size of both types of incentives taken together, since stronger incentives from either side are expected to have more effect. The second component of the index is a measure of the extent to which positive incentives are larger than negative incentives. A high index value will indicate that the incentives are overall large, and mostly positive. A low index would then indicate low overall incentives,

which are mostly negative. Combinations of either low and mostly positive incentives, or high and mostly negative incentives will result in average index values. Appendix A explains in further details how this index is measured.

The index values are illustrated in figure 11.4, where they are plotted on the x-axis. With the impact being plotted on the y-axis, the relationship between the incentive structure in the program and the outflow to work becomes visible. Clearly, this relationship is a positive one when looking at the general trend³⁴. This is even more so when considering the fact that when correcting for target group in the ND25+ and the Temporary Job program, these two programs would be found slightly higher in the graph. Although this figure does not fully explain variation in impacts (for example the sharp jump between Ontario Works and Work for the Dole), the ranking with respect to the index does indicate a correlation between the impacts and the carrots-and-sticks index.

Figure 11.4 The carrots-and-sticks index and the impact of the programs



The implication of the correlation between the carrots-and-sticks index and the outflow to work ranks is significant. What can be concluded is that through focusing on negative incentives alone, programs will undermine their effectiveness. The number of programs in the benchmark is not

³⁴ The correlation coefficient for the carrots-and-sticks index and the outflow to work ranks is 0.78. A regression analysis between these two variables shows that the index is statistically significant at the 5% level and that the R^2 is 0,60 (and the adjusted R^2 is 0,53).

large enough to perform more sophisticated analysis on single indicators, but some careful considerations are possible.

With respect to sanctions and activation conditions as part of the input of the program, these were lowest in the United Kingdom and in the Netherlands. These low ranks on these two indicators explain how the Employment Option of the NDYP had reached very high level of input-efficiency, since it these low scores meant that input was small while results were high. When looking specifically at the NDYP Employment Option and the Work First programs, it can thus be said that lower sanctions and activation conditions did not undermine good results. This of course has to be seen in a relative context, since these can be considered nevertheless high in an absolute sense.

In addition, low benefits, defined here as a high level of negative incentive, do not seem to affect outflow to work in the UK, since performance was the highest in the NDYP Employment Option. However, the combination of low benefits with high rewards might be the key determinant here. On the other side, the low rewards in the Dutch Work First projects does not affect its rate of outflow to work, but the benefit level was higher than in the other programs. Here again, the importance of balancing “sticks” with an appropriate level of “carrots” seems to be important for good results. This is confirmed by the fact that Work for the Dole and Ontario works combined low benefits with low rewards. Nevertheless, high benefit levels and relatively high rewards in the two Swiss programs where not influential enough to guarantee a high level of impact in these program.

The ranking of the indicator for the intensity of the program was based in the fact that a high number of hours would be expected to lead to better outflow to work, and thus high intensity received a high rank. This thus means that a low number of hours in the program would result in a smaller level overall level for the sticks in the program. It was however the case that low intensity was mostly found in Ontario Works, Work for the Dole, and RMCAS, three programs with actually high levels of negative incentives. This outcome points towards the fact that programs with a stronger focus on negative incentives do not require many hours of work from their participants, possibly because the high focus on negative incentives could not justify such a choice.

In conclusion, synergies between different elements in work-based employment programs are influential. For optimal outflow to work, a balance in positive and negative incentives needs to be present. As just discussed above, programs who focused only on negative incentives

where found to be inefficient. Moreover, low levels of sanctions and activation conditions, two important “stick” elements of work-based programs, were much less important in the best performing programs. The reason for this could also be that the relationship between these indicators is not linear. This would imply that at some point, more stringent activation conditions and harsher sanctions would not automatically lead to better results. It is however not possible to firmly test such a hypothesis in this benchmark, as more data-points would be needed. Nevertheless, the section next will look further into the way sanctions are related to impacts, by linking sanction laws to their procedure and the actual number of sanctions used.

At last, attention should be given to the different indicators which make up the positive incentives of the index, as these are proven to be important determinants of success for the programs. Combining a private sector work-environment with regular wages in a program lasting 6 months and providing a sufficient amount of training and job search assistance seem to be the reason why the Employment Option was able to reach such high level of outflow to work. The same can be said for the Dutch Work First programs in this benchmark (see Sol, Castonguay, Van Lindert and Van Amstel, (forthcoming) for more information on the importance of these determinants of success within the various Work First projects in the Netherlands).

11.4. Sanctions: laws, procedures and quantities

Sanctions are a fundamental element of work-based employment programs, since they make it possible to enforce the mandatory character of the program. Nevertheless, the previous section already showed that “softer” sanctions do not harm the rate of outflow to work and that harsher sanctions do not guarantee better outflow to work. Of course, softer has to be interpreted in a relative way, as the ranking is based on what is found in the other programs in the benchmark. Interestingly, sanctions are not only part of the input of the program, but are also part of the process and the output. Within the process, sanction procedures were measured for the extent to which these are flexible and formal. Following this, the output measured how many sanctions are actually being used within each program.

Many hypotheses can be made concerning the links between these three components. The first one is that harsh sanctions at the input level will indicate a commitment to strict monitoring, leading to the design of a

flexible and formal sanctioning procedure in order to properly apply those sanctions, leading to a large number of sanctions being used due to the strong focus on sanctioning as well as a fair and clear system for applying sanctions. Of course this could go the other way around: harsh sanctions would be paired with inflexible and informal procedures and leading to few sanctions being applied because of the reluctance of the case-managers to use them. Many other combinations are in fact possible.

What can be seen from this benchmark are four models with respect to the combination of sanction legislation in the input, sanction procedure in the process, and sanctions quantities in the output. These three indicators are summarised in table 11.2 below. As was explained in the output-benchmark, data collection with respect to the quantity of sanctions used within each program is very poor in most programs. Precise comparisons are thus difficult to make, but based on the quantitative and qualitative evidence discussed in chapter 8, a distinction between lower quantities and higher quantities is made in the table.

Table 11.2 *Sanctions throughout the policy-chain*

	Input Legislation	Process Procedure	Output Quantities
United Kingdom <i>NDYP</i> <i>ND25+</i>	Softer Softer	Formal/ Inflexible Formal / Inflexible	Higher Higher
The Netherlands <i>Work First</i>	Average	Average formal / Flexible	Higher
Australia <i>Work for the Dole</i>	Average	Average formal / Average flexible	Higher
Canada (Ontario) <i>Ontario Works</i>	Harsher	Informal / Flexible	Higher
Switzerland <i>Temporary Job</i> <i>RMCAS</i>	Harshest Harsher	Formal / Inflexible Informal / Inflexible	Lower Lower

On one side are the two programs in Geneva, where an “all-or-nothing” approach is found in the legislation, and where little importance is given to sanctioning procedures within the design of the programs. High flexibility and low formality has thus meant that case-managers where given a lot of freedom in interpreting sanction requirements. Especially in the case of the RMCAS, evaluations of the program specify how interpretations vary greatly from one case-manager to the next, but that in general the strictness of the sanctions has been diminished by many (see

Cunha et. al., 2002, p.33). This has lead to a lot less sanctions being used in this program, through a clear reluctance of completely removing the benefit of individuals who did not meet their activation requirements. Sanctions were arranged through more formal procedures within the Temporary Jobs programs since participants actually fell under employment protection regulations. No data was officially available on the quantities of sanctions used, but analysis of the information available indicates that this number is expected to be low.

The second category is made up of the Ontario Works program, where sanctions are central to the programs' approach. This has lead to harsh sanctions being introduced in the legislation, which are applied in an informal setting by case-managers. Sanctions are nevertheless allowed to be adapted to the seriousness of the breach, and where thus described as flexible. This has lead to many sanctions being used by case-managers, as reported by surveys by Lightman, Mitchell and Herd (2003 and 2006). Making a link with evidence on the effectiveness of these programs, it is obvious that these two models for sanctions did not lead to high rates of return to the labour market.

The Dutch Work First projects and Work for the Dole present relatively similar patterns with respect to sanction strictness, procedure and use. Average levels of sanction impacts on benefit income where applied in a relatively formal manner, and in the case of the Work First projects, with relatively higher flexibility. These input and process produced a relatively high amount of sanctions, with one in four being sanctioned either for refusing to participate or through misconduct or leaving prematurely. However, it was mentioned how this high quantity of sanctions in the Netherlands is likely to be overestimated, such that one can expect that the number of sanctions would be somewhere in between the high and the low values. All in all, these programs' performance level was also found in between the best and the worse rankings, leading to the conclusions that average levels of sanction-elements in the policy-chain lead to average outflow to work.

The two New Deal programs had the less harsh sanctions in this benchmark, but also the most formal and inflexible sanctioning procedures. Hence, case-managers did not have much freedom in interpreting the sanctioning legislation, and sanctions where used relatively frequently. Data is not available for the work-activities of the ND25+, but since it has similar overall number of sanctions used within the whole program as the NDYP, it can be expected to have also similar numbers with respect to its work-based phase. This combination of a formal sanctioning procedure, which could also not easily be adjusted

from case-to-case by the case-managers can be interpreted as a procedure in which some sanctions might be deemed unfair because of the less appropriate match between the action and the consequences. But on the overall level, since formal procedure leads to a much clearer set of rules, allowing claimants to better calculate the effect of a breach on their income, also means that sanctions might be better predicted and expected by claimants. Combined with lower effect on total income, such sanctions might have a lot more power in correcting uncooperative behaviour, by providing a “wake-up call” without having too harsh financial consequences. Furthermore, it should be noted that sanctions rely on the rational calculation of claimants of the impact such a sanctions can have on their income, hoping to deter them from choosing not to participate or not to properly look for jobs. Within this perspective, formality in the procedure is thus crucial for the proper working of the incentives structure intended by sanctions. Results varied between the different New Deal programs, such that a clear line between this sanctioning model and the results it reaches is difficult to make. Nevertheless, considering the target group and the process within the ND25+ and the NDYP VS/ETF, it can be said that the sanctioning model of the New Deal programs did lead to relatively good results, especially in the case of the Employment Option.

11.5. Conclusions: the bright and dark sides of work-based employment programs

This explorative benchmark of the design and the results of work-based employment programs has two main conclusions. The first one is that work-based employment programs all vary greatly from each other. They vary not only in their approach to solving unemployment, but also in the result they are able to achieve. As a result, when discussing the success of work-based employment programs and their approach to reducing unemployment, attention should be paid to which specific type of program is being looked at.

Indeed, this benchmark has shown that very successful programs are found in the United Kingdom and the Netherlands. A closer look at the inputs, process, and output of these programs has lead to the identification of important determinants of success. On the other side the programs in Ontario and Geneva where not able to assist many of their participants in finding jobs, and the Australian program performed in between those two groups of programs, together with some other programs within the UK. Analysing the policy-chain of these programs

allowed drawing some lessons from which all programs can learn. This benchmark has thus shown that work-based employment programs have both a dark and a bright side, as shown by the sample of programs in this benchmark.

The second conclusion that was reached by this benchmarking exercise was that choices in input, process and output mattered greatly for performance. One important explanatory factor for high impacts was that input, process and output altogether need to be optimal. Low levels in all three elements of the policy-chain naturally resulted in low impacts, but most interesting is the finding that in-between combinations were also ineffective. The second key explanatory factor for high outflow to work was the importance of positive incentives in the programs. While all programs had rather high levels of “sticks” or the negative incentives, the most effective programs balanced these with even higher levels of positive incentives. Thirdly, with respect to sanctions, this it was concluded that a low level and duration of sanctions did not undermine outflow to work, as long as these were matched by formal sanctioning process in which the effect of breaches on income is easily predictable.

By designing work-based employment programs according to these determinants of success, the bright side of “workfare” is thus put on the foreground. Instead of focusing on threatening the unemployed into jobs, work-based employment will reach better results if their approach is centred on all three elements of the intervention-strategy. Increasing hard and soft skills, improving motivation to take part in the labour market, and opening access to opportunities in the labour market will lead to better results if synergies are being created between these elements. Such an integral intervention-strategy is also more likely to lead to more sustainable jobs with decent wages, as employability is enhanced from different angles. Such an approach is also encouraged by the OECD (2003, p.213) as it explains how the combination of strict job search requirements without the provision of job search assistance and guidance can result in entry into precarious jobs, but that job search requirements matched by intensive supervision and assistance may improve employment earnings. Work-based employment programs, especially under the name of “workfare”, are usually assumed to be actually very unbalanced with respect to their intervention-strategy which is expected to be mostly relying on negative incentives. This benchmark proves that this does not need to be the case. Even more so, the most important conclusion reached by this benchmark is that by focusing on negative incentives alone, outflow to work is actually undermined. Investments in appropriate levels of positive incentives in the form of provisions which

are able to support the participants in their return to the labour market are thus shown to be efficient, since overall results are then improved.

From the perspective of the demand-side of the labour market, work-based employment programs have the potential to become an important instrument for making sure the labour market programs can accommodate the needs of employers. Work-based employment programs can have a large impact on access to work through providing a relevant work experience for their participants. In fact, determinants of success such as a private working environment, a salary and a substantial number of hours per week in the program will all contribute to making the work-activities within the programs a relevant work-experience in the eyes of potential employers. Moreover, the possibility that employers retain the program participants after the end of the program is also one important reason why the Employment Option is so successful in assisting its participants in finding a job. Labour market programs have in general also attempted to increase access to work by providing assistance with matching the unemployed to the right job offers and assisting in the application process for these jobs, such as helping out with writing a résumé. It should thus be clear that work-based employment programs have an important value-added with respect to traditional job search assistance, since they can also have a direct effect on what the unemployed can write in their résumé in terms of work-experience.