



Governance and Performance of Public Employment Services in the United States and the European Union: Lessons and issues for discussion

3 – 4 September 2015

Globalisation and demographic changes have led to similar labour market developments, and have posed similar challenges, on both sides of the Atlantic. Amid rising long term unemployment and concerns over fiscal sustainability, governments are under growing pressure to improve the efficiency of public employment services (PES). This has increased interest in the methods for measuring PES performance and the effectiveness of active labour market policies.

Although they are facing similar challenges, the US and EU labour market institutions differ in several respects, for example in the level of minimum wages, the role of unions, the level of spending on active labour market policies, or the flexibility of employment protection legislation. There are also large variations across EU member states¹, for example between countries with high levels of labour market regulation and a strong PES offering a wide range of services (as in Germany) and those with a relatively limited provision of ALMP and activation services (as in Spain). This implies that a rich variety of practice and experience is available both in the US and in the EU to offer useful insights on how to improve performance in a given institutional context.

The workshop was initiated by the Atlantic Council in cooperation with the University of Maryland's School of Public Policy and organised with the European Commission. The discussion focused on PES programmes in the European Union and the United States, the challenges and opportunities they faced and the approaches followed to monitor and evaluate the performance of labour market policies. Participants included government officials, researchers and senior PES managers from the US and the EU.

This brief report aims to highlight the key conclusions and issues that may merit further discussion, rather than provide a complete account of the rich discussions that took place at the workshop. The lessons are grouped around the following four issues:

- Improving the effectiveness of PES measures and services, and target setting;
- Measuring performance and identifying high performers;
- Identifying efficient practices; and
- Promoting the sharing of good practices.

1. Improving the effectiveness of PES measures and services, and target setting

The discussion on governance systems, accountability and target setting yielded some interesting insights about how to improve PES performance through institutional arrangements, performance incentives or contracting out.

1.1. Decentralisation

The decentralisation of managerial functions, such as target setting, policy design, or the management of partner relations may be one way to improve PES performance. The US has for many

¹ For example, spending on active labour market policies ranged between 0.22% in Slovakia and 1.82% in Denmark according to OECD statistics for 2013.

years operated a decentralised system of decision making and service delivery. Reflecting the broader developments of regionalisation in public government and the extended use of New Public Management tools in public administration, several EU member states have also moved towards decentralisation in administering PES programmes (EC 2011). Decentralisation may increase accountability and foster cooperation with stakeholders at the local level. It may also allow more room to adjust policies to local needs. In fact, the combination of decentralisation with a formula based block grant can be used as a performance management tool.²

Apart from these widely accepted potential advantages, workshop participants stressed the importance of channelling in more information from local managers about local labour markets and actors, and also about their own resources. This resource can be tapped, for example in bottom-up planning procedures and especially target setting, in order to establish realistic targets that account for local constraints and resources (as for example in Germany).

A lesson that emerged from a recent OECD study is that allowing a high degree of flexibility in programme management at the local level tends to result in improved PES performance (OECD 2011). A good example of this is the waiver system introduced by the US Federal Department of Labor, which allowed states to apply for additional flexibility when implementing innovative workforce strategies and initiatives (OECD 2011). A similar system is used in the UK: the so called '[City deals](#)' allow cities to negotiate for more freedom in how they use their resources to meet local needs and manage development projects. In Austria, the PES target setting procedure allows for some flexibility: local PES units can obtain additional funding if they set more ambitious targets. The variety of examples suggests that there is some room for flexibility in any institutional setup (OECD 2011), however, limited resources or political instability may constrain the actual use of flexibility even where the legal framework allows for it (as for example in Poland or Spain).

While decentralisation can be effective, it must be designed with care: some functions can be decentralised while others should be kept at the centre. For example, in the Danish system, the centre maintains a freely accessible data system and the platform of e-services for all job-seekers, which ensures transparency and lowers the overall costs. In Finland, a standardised protocol for PES services has been defined at the central level in order to ensure the same level of quality in PES services across local units. The definition of monitoring and performance indicators, and data collection methods, are also important to be centrally provided in order to ensure data comparability.

Decentralised service provision may also give rise to high regional variations in the quality of, or access to, services. The risk of such distortions can be mitigated by the close monitoring of accessibility and performance, and also by strong political accountability. For example, the Danish system relies on both mechanisms, and also maintains a highly transparent information system, which reinforces both managerial and political incentives.

Exploiting the potential benefits of decentralisation requires considerable analytical and management capacities at the local level. Although this may give rise to concerns over the availability of sufficient managerial expertise at the local level, there are several examples to prove that it is possible to develop the local capacities necessary for a decentralised system (OECD 2011).

² For example, in the Dutch reform of 2004, municipalities received an annual grant from central government to cover the expected - rather than the actual - cost of social benefits, which is calculated on the basis of local demographic and labour market characteristics. This represents a benchmark for expected performance, and a strong incentive for the municipalities to keep the number of benefit recipients at the expected level (or below) by maintaining effective and efficient reintegration services. ([Danish Ministry 2013](#)).

Issues for further discussion:

- How to strengthen horizontal accountability (e.g. lessons from four-party associations in Korea, role for local democracy)?
- How to ensure coherence between national goals and local level targets?
- How to handle the mismatch in the boundaries of local labour markets and administrative units?
- What lessons may be drawn from a review of block grant formulae used in decentralised systems?

1.2. Managerial incentives

Although performance management techniques are widely used by PES in the EU, financial incentives for managers are not yet commonly applied. The Austrian and German examples seem to have yielded good results in terms of improved PES performance, although it is hard to separate their effect from the other elements of their incentive systems. In Denmark, financial incentives are believed to have led to more peer-learning across PES than by using the open method of coordination system on its own.

An obvious precondition for financial incentives to work well is that they are tied to performance indicators that managers have some control over. For example, where wages are more difficult to influence than placements (as seems the case in US labour markets), wage outcomes should have less of an impact on rewards for managers. The definition and measurement of indicators and the adjustments for the local context should be transparent and fair, and accepted as such by managers.

However, even otherwise well-designed financial incentives for managers or PES units can have unintended side effects, such as a decline in risk taking or the lack of a long term vision. Competition over budgets may backfire, as it makes it difficult to plan ahead.

Beside financial rewards, recognition can also be a strong incentive. In Austria for example, the highly publicised ranking based on the balanced score card³ scores and the annual award ceremony to celebrate high achievers may even be more important than the managerial bonus system. Another non-monetary tool that has been used in the US is to reward high-achievers with more flexibility in the target negotiation procedure for local service providers.

Benchmarks and feedback on performance can also motivate managers⁴, and help support performance improvement in a decentralised system. Cluster based benchmarking is expressly used in both Austria and Germany as part of 'bottom-up' target setting procedures, inviting managers to look at the indicators of the best performers within their cluster before they set their goals. A similar practice has been developed by the W.E. Upjohn Institute for Employment Research, in which the benchmarks come from statistical analysis rather than peer PES: they forecast potential outcomes based on the personal characteristics of clients, so that providers can compare their actual performance to this benchmark.⁵ Other forms of feedback, such as qualitative assessments and guidance on possible improvements, are also widely applied. The regular country specific

³ The BSC is a strategic management tool that objectively compares PES organisational units against each other on the key principles of AMS' mission statement. The BSC is weighted to take account of key resource aspects including staffing and budgets and different local situations. It helps to define targets and focal points for development for the individual PES offices and, on a wider level, it helps to implement the PES organisation's overall mission strategies.

⁴ Feedback from customers was also mentioned as an important driver that has been influential in shaping recent reforms in VDAB, Belgium.

⁵ A similar estimation is used in the UK to set the minimum performance level for external providers.

recommendations of the European semester, and more specifically the recently established benchmarking system for the PES, are examples of such approaches.

Lastly, stability itself (of both funding and the manager's position) can contribute to high performance by enabling PES leaders to plan ahead and undertake reforms that only yield returns in the long term. Stability can be granted by giving a strong mandate to social partners, as in Austria where the PES leader is appointed by a tripartite board and the government cannot unilaterally intervene. Another way to reduce politically-induced volatility is to tie funding to indicators of need, as in the case of the formula funded programmes of the US.

Issues for further discussion:

- How to foster ownership of targets among managers and staff (to avoid hitting the target but missing the point?)
- How to encourage risk taking in a bottom-up target setting procedure?
- How to introduce strong managerial incentives in a context of strong employment protection legislation and powerful unions?
- How to combine quantitative and qualitative indicators in performance management?

Target setting and management incentives in Germany

In the German system national level targets and objectives are set in contract-type agreements, in a bottom-up process involving the local PES branches. Past performance and expected outcomes within the benchmarking cluster are used as a basis for the negotiations. The national level however retains the power to adjust targets. For example, they lowered outcome targets when growth forecasts signalled a significant slowdown at the beginning of the global financial crisis.

Performance dialogues are closely linked to the target setting process and represent the main means of performance management. Dialogues are held quarterly at all hierarchical levels – between the board of directors of the head office and the ten regional directors, as well as between the ten regional directors and respectively the local directors of the 156 local agencies.

The Bundesagentur für Arbeit, the German PES, considers this holistic approach essential for effective performance management, which enables the organisation to evaluate performance in the context of the local situation and capabilities. Performance dialogues provide an open space to present the challenges, discuss and agree on the targets and interpret the results at the respective levels. Performance dialogues apply the main principles of the Common Assessment Framework (CAF).

The central, regional and local controlling departments take a lead role in preparing the events. They are responsible for proposing the main issues to be discussed, presenting the necessary background information and moderating the debates, as well as for reporting and briefing.

1.3. Contracting out

The contracting out of PES services can be a way to increase overall performance via introducing competition and/or customer choice. It can also increase the flexibility of service provision, and support the fast reallocation or expansion in capacities or infrastructure. However, it is not very wide-spread in Europe, and is mainly used in training programmes or as way to serve clients that require specialised expertise or to increase PES capacities (Fuller 2011). Somewhat surprisingly, there is no discernible time trend in the use of subcontracting. Payment by results arrangements are common in the US, but relatively rare in Europe, where PES typically use fixed price contracts. In Europe, the existing experience in subcontracting is thus limited and also somewhat mixed. The

Dutch reform that introduced results-based subcontracting in 2004 has received much criticism, though mainly on political grounds.⁶ In the UK, private providers involved in the recently introduced large contracts have gradually improved their performance and many have superseded the expected placement targets, although a recent report by the National Audit Office has noted the lack of credible evidence on the cost-effectiveness of such contracting arrangements.

In the UK, payment by results contracts have been introduced with a soft start: initially providers were entitled to an attachment fee for referrals (not only placements), which was gradually phased out. The UK 'Work Programme' was launched in 2011 and included 18 'contract package areas', covering 40 contracts between providers and the Department for Work and Pensions (DWP) running until 2017. Payment is tied to performance over the expected minimum placement rates, which are calculated separately for specific client groups, based on the estimated probability of re-employment (considering their age, education and other relevant characteristics). DWP closely monitors providers to suppress gaming, and regularly audits the accuracy of reported performance figures. Based on monitoring information and monthly performance reviews, they regularly adjust the terms of the contracts. While the providers have performed above expectations, the contract management costs have also been substantial⁷ so the net gains of the programme are small.

Issues for further discussion:

- How to ensure that contractors are not exposed to excessive risks while avoiding excessive profit margins? ?
- How to handle political concerns about contracting out?
- How to protect providers from weathering economic downturns (if at all)?

Payment by results approaches in the US and the UK

Payment by results methods have been widely used in the United States and the UK. The Work Programme is the most significant scheme of this nature in the UK, which involves large contracts and a payment by results approach to incentivise providers to secure improved, and sustained, outcomes for their clients. An associated development is the increased use of social finance in labour market programmes. One example is the use of [social impact bonds or SIBs](#). In a SIB, a financial intermediary raises private capital that is used to deliver services that ordinarily would be provided by the government with the government providing a payment (with interest) for successful outcomes. Developed in part to enable charities and third sector providers deliver payment by results contracts, SIBs are an innovative opportunity to access external finance for the delivery of services.

Different models of SIBs have emerged across the UK, which include a range of models and can feature intermediary bodies to broker funding arrangements between social investors and providers. The principal innovation of a SIB, in conjunction with the payment by results contract, is that the risk associated with non-performance is transferred to the investors instead of the taxpayer.

A second example is the 'pay for success' model that has also gained some popularity in the US. In a pay-for-success contracting arrangement, the government makes payments to private service providers only when the desired outcomes are achieved. Though the recent recession has reduced funding sources, there are about 10 projects running at the moment, some of which are implemented by [Social Finance US](#). The existing [evidence](#) on the efficiency of these approaches tends

⁶ For more detail on the development of the Dutch system see [European Commission \(2011\)](#).

⁷ No estimate is available on management costs in the UK. Based on the Australian practice, management costs may amount to 6 % of total spending.



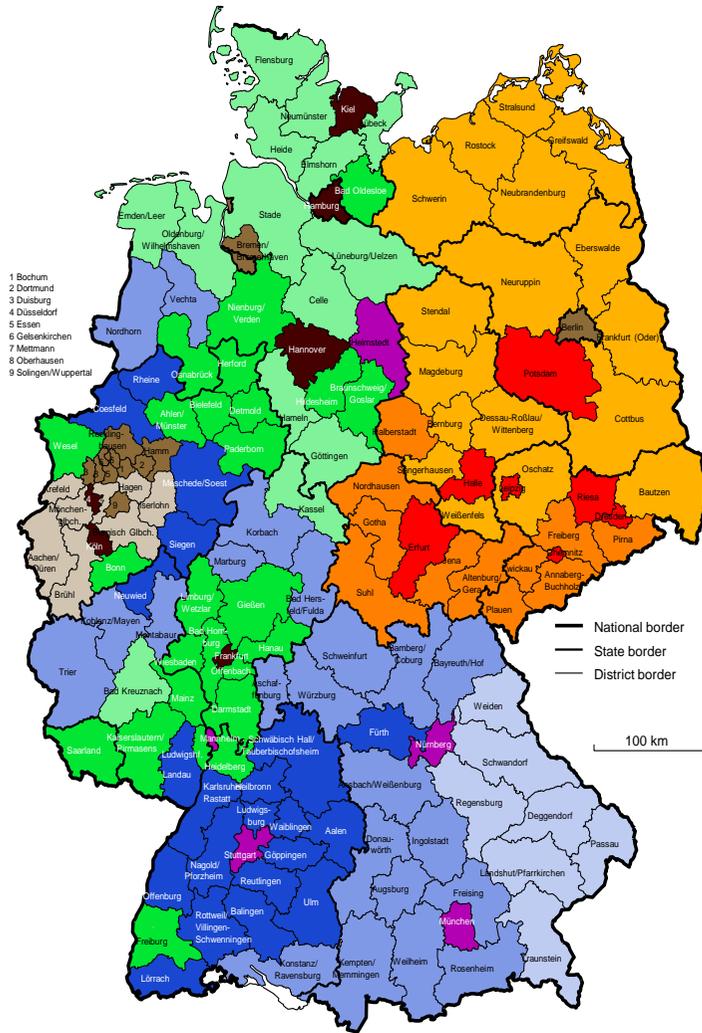
to be mixed, but experts agree that there is great potential in them to raise expand the available funding and increase efficiency.

2. Measuring performance and identifying high performers

The incentive systems outlined in the first chapter all necessitate the accurate measurement of performance. The discussion on benchmarking systems brought forth some common challenges in defining indicators and collecting data.

On both sides of the Atlantic, governments and the PES use various methods to account for job-seeker characteristics and the local context in performance indicators and benchmarking systems, which may range from calculating outcomes by sub-groups to regression based adjustments and clustering. In the US, the regression approach is the most wide-spread way of controlling for local factors, while EU countries (notably, Austria, Germany, and the UK) tend to use clustering (or simpler methods). The regression approach produces a ranking on a single scale, while clustering has the advantage of defining groups that share similar characteristics. For example, in the case of Germany, clusters are characterised by the share of the urban population, the unemployment level, and also by the seasonality in unemployment, induced for example by tourism (and a few other control variables). As the figure on the following page shows, the current system groups local PES units into 12 main clusters. Members of the same clusters are often in the same geographical area, but in some cases (for example metropolitan areas) they are spread across the country.

Figure 1. Clusters in Germany.



- Type I (5): Predominantly metropolitan districts with favourable labour market conditions
- Type IIa (6): Metropolitan districts with above average unemployment rates
- Type IIb (11): Metropolitan districts with very high unemployment rates
- Type IIc (8): Urbanised districts with slightly above average unemployment rates
- Type IIIa (25): Districts with conurbational features with below average unemployment rates
- Type IIIb (14): Predominantly rural districts with average unemployment rates
- Type IVa (21): Districts with conurbational features with a large manufacturing sector and favourable labour market conditions
- Type IVb (22): Predominantly rural districts with favourable labour market conditions and strong seasonal dynamics
- Type IVc (7): Rural districts with very strong seasonal dynamics and low unemployment rates
- Type Va (7): Predominantly metropolitan districts with high unemployment rates
- Type Vb (11): Predominantly rural districts with high unemployment rates
- Type Vc (17): Rural districts with very severe labour market conditions

() Number of districts in each type

Source: Statistics department of the Federal Employment Agency

© IAB 2013

Obtaining reliable and accurate data on performance seems to be a challenge in both the EU and the US. Some of the difficulty arises from institutional features, for example, that external providers may lack the necessary IT infrastructure. Reliable and comparable data also presupposes some degree of standardisation, especially in the case of process indicators: if there are significant variations in the protocol of a service (or if there is no protocol at all) comparison across providers is meaningless. Further distortion can arise from the motivation of actors within the system, especially if financing is tied to the indicators. Subcontracting and, more generally, the fragmentation of the service market (which seems especially relevant in the US) can increase these difficulties.

Efficiency indicators appear particularly difficult to generate, mainly due to the difficulty of identifying the cost of PES inputs per intervention and per client when the services are provided internally by the PES. In the US this is further complicated by the use of multiple funding streams.

Data quality can be improved if data collection and analysis is designed in such a way that the service provider providing the data can also benefit from the resulting database: for example if the provider uses the monitoring data to detect problems in service design or delivery. A good example of this comes from the W.E. Upjohn Institute for Employment Research who use data from the unemployment register to generate a baseline estimate for the expected labour market outcomes of a group of clients, which is shared with the provider to be used as benchmark for their performance.

Another option is to make funding conditional on collecting and supplying reliable monitoring data, which has been recently introduced for EU-funded projects under the [European Social Fund](#) programme. The administrative cost of data collection can be reduced (especially at the national level) by providing a standardised IT platform for data collection freely available to all types of providers, as for example in the Danish system.

Recent developments in IT and administration also offer new opportunities for improving the information base. For example, employment outcomes can now be measured more easily by linking the unemployment register to tax or insurance records on employment. This method may make it possible to observe long term outcomes at very little additional cost. Another option to contain costs is to use existing survey data as a baseline estimate for potential outcomes.

Issues for further discussion:

- What should be considered as outcomes beyond placements (e.g. job quality, skills upgrade, wages)?
- How can the PES handle the tension between needing more time for obtaining good outcome measures and the time service providers must wait for their fees?
- Can clustering lead to perverse incentives (e.g. for high performers within a poor cluster)?
- How to tackle barriers to linking administrative data (e.g. personal data protection rules)?

3. Identifying effective practices

While performance measurement can identify high achieving *units or individuals*, impact evaluation can detect which *practices* are the most effective.

The identification of effective measures is often based on a comparison across programmes and/or providers. However, evaluation results and outcome indicators for monitoring purposes are not always comparable across measures, for several reasons. For example, short term indicators

disfavour measures that only yield results in the long term (as is typically the case in training programmes), and contextual factors may also differ considerably.

Cross country comparisons of similar measures may be more complex than within country ones, due to unobserved variations in the local context, or in the interpretation of the indicators. Impact evaluations often measure the impact of a programme as a 'black box', without specifying what aspect or element of the programme makes it work well. However, there are several ways to identify such aspects. Qualitative techniques can analyse the programme design and implementation based on interviews, and logic models can explore the possible causal relationships between programme elements and outcomes using economic, psychological or other theoretical models. These are routinely used by the US Department of Labor to identify good practices. Quantitative techniques exploit reform episodes or planned experiments to identify the impact of changes in programme design (or implementation). The latter typically involves a randomised controlled trial⁸ in which the baseline programme (A) is tested against a slightly modified version (B), and is therefore often referred to as A/B testing by US practitioners. These are widely used in the US, and also in some EU countries, especially in Denmark, the Netherlands, Germany, and the UK.

Participants agreed however that there is a need for more and better evaluations. Centrally provided training courses and legal requirements or conditionalities in the funding are used in both the EU and the US to promote evaluations, but this may not be sufficient to promote a stronger evaluation culture. There are some promising examples of capacity building in using simple evaluation methods that support the refinement of services, e.g. in the W.E. Upjohn Institute for Employment Research. Apart from generating useful evidence, the involvement in such in-house evaluations can also increase the engagement and motivation of service providers. In the EU, the [Centre for Research on Impact Evaluation](#) has launched a series of pilots to carry out counterfactual impact evaluations of projects financed by the European Social Fund.

Issues for further discussion:

- How to mitigate the potential ethical concerns about randomised control trials (A/B testing), e.g. excluding a client from a service element they may benefit from?
- What are the main barriers to developing a strong evaluation culture?
- How to tackle these barriers?

4. Promoting the sharing of good practices

Once effective practice has been identified, how can the resulting learning be shared to drive improved PES performance? The discussions investigated a series of approaches including 'benchlearning' and peer learning.

The EU has just established a more formal network for the PES, which provides a framework for a novel approach called 'benchlearning'. The idea is to create a systematic link between benchmarking and mutual learning activities by identifying what works and why using quantitative and qualitative assessments, and developing targeted mutual learning activities between PES on this basis.

The focus of the network is on sharing experiences, and no financial incentives or sanctions are involved. Peer learning is supported by a series of activities, including thematic reviews of PES practices, study visits, mutual assistance projects, toolkits and studies.

⁸ This is a research design in which participants are randomly allocated into two (or several) groups, and each group receives a different programme (treatment). The random allocation ensures that the groups have the same characteristics (on average), and thus, cross-group variation in outcomes will only be due to differences in the programmes.



Clustering is used in some EU countries as a way to support peer learning. Best performers in each cluster can be relevant role models for the other units within the cluster as they operate within a similar context. Austria also uses a financial incentive: they reward local PES for transferring good practices. Interestingly, the bonus is shared between the contributing and receiving party.

The US government does not sponsor such a formal approach to information dissemination; meetings of professional associations and academic conferences are the primary vectors for peer learning.

Issues for further discussion:

- What are the most efficient formats (seminars, study visits, repositories etc.) for exchanging detailed information on good practices?
- What are the main barriers to sharing good practices across PES, and how can these barriers be addressed?
- Practically, what steps must be taken to enable PES to adapt good practices to meet their own needs?



References

Danish Ministry (2013) Decentralization with national standards. The case of the Netherlands, Maarten A. <http://oim.dk/publikationer/c/kap06.htm>

European Commission (2011a) Decentralisation of Public Employment Services. Analytical Paper. PES to PES Dialogue, The European Commission Mutual Learning Programme for Public Employment Services. DG Employment, Social Affairs and Inclusion. Author: Hugh Mosley.

European Commission (2011), [Subcontracting in Public Employment Services: Review of Research Findings and Literature on Recent Trends and Business Models](#), DG Employment, Social Affairs and Inclusion, The European Commission Mutual Learning Programme for Public Employment Services. [Britain, Germany and the Netherlands] Author: Dann Finn.

Fuller, Andy (2011) (ed) Partnership among employment services. Small-scale study. Ismeri Europa/Alphametrics. <http://ec.europa.eu/social/BlobServlet?docId=7122&langId=en>

JSA (2015) Job Services Australia Star Ratings Methodology from July 2012 to June 2015, Job Services Australia <http://docs.employment.gov.au/system/files/doc/other/jsa-star-ratings-methodology-july-2012-to-june-2015-fact-sheet.pdf>

NAO (2014) Cabinet Office. Transforming government's contract management, National Audit Office, UK, September 2014 <http://www.nao.org.uk/wp-content/uploads/2014/09/Cabinet-office-cross-government-transforming-governments-contract-management.pdf>

OECD (2005) Public Employment Services: Managing Performance, in: Employment Outlook, 2005.

OECD (2011) Building Flexibility and Accountability Into Local Employment Services: Synthesis of OECD Studies in Belgium, Canada, Denmark and the Netherlands Francesca Froy, Sylvain Giguère, Lucy Pyne, Donna E. Wood <http://www.oecd-ilibrary.org/docserver/download/5kg3mkv3tr21.pdf?expires=1442255161&id=id&accname=guest&checksum=258BD1A3E9FAFA04AC2FF6E4913FB42C>