European Platform Undeclared Work

Follow-up Visit on Risk Assessment for More Efficient Inspections

28 September 2018, Brussels, Belgium

Summary Report

1.1 Introduction

This report summarises the information shared by the Belgian National Social Security Office (NSSO) with members of the European Platform Tackling Undeclared Work (the Platform) during a visit to their premises on 28 September 2018.

The follow-up visit was part of a mutual learning process, related to a Thematic Review Workshop (TRW) on 'Risk Assessment for More Efficient Inspections[1] which took place in Madrid on 14-15 June 2018. During that workshop, the participants nominated Belgium as one of the countries with a well-developed risk assessment system, that would merit a follow-up visit. Belgium offered to host the visit to discuss in-depth the country's risk assessment strategy, system and operations.

Representatives from ten EU Member States (Cyprus, Czech Republic, Germany, Ireland, Latvia, Lithuania, Portugal, Romania, Spain, Sweden) participated in the visit, as well as representatives from the European Commission (Platform Secretariat).

1.2 Framework of the host organisation

Key messages for policymakers:

Developing a risk assessment system is an ongoing process that needs support at many levels - including 'champions' at top government level and willingness to invest in both skills and tools for risk assessment.

Branding a database with a memorable name, and providing additional support to users, can help to increase the quality and quantity of data entered to a system.

The Belgian Social Security Organisation (NSSO) has developed seven key e-government database projects since 1997. It took two decades to set up this risk assessment system which required long-term senior governmental support. The databases were accessible to various institutions and inspectorates, and provided real-time consultation and 'big data' for the use of datamining. These include among others:

- DIMONA (mandatory declaration of the start or end of a labour relationship);
- LIMOSA (mandatory declaration of posted workers and self-employed);
- A database for the declaration of building sites and subcontractors (real estate and building sector with a contract value of at least EUR 5 000 with at least one sub-contractor or with a contract value of at least EUR 30 000 with or without a sub-contractor);
- DMFA (multifunctional quarterly declaration of wages, working time etc.);
- Checkin@work (the daily registration of workers in the real estate and building sector).

[1] The report from the Thematic Review Workshop can be accessed here http://ec.europa.eu/social/BlobServlet?docId=19924&langId=en

The social security institutions and their databases are interconnected via the 'Crossroad-bank for Social Security (CBSS)', a public body. Via the CBSS, the main agencies tackling undeclared work in Belgium have access on a permanent basis to these databases. Sharing of data with other relevant ministries and institutions should be a priority and when setting up registers or using databases or data collection, special care should be given to ensure that the levels of confidentiality assigned to these data do not impede the sharing of information between relevant authorities.

Belgium has found that it works well to have an easy to remember 'branding' for the different databases to raise awareness of the systems and to prevent problems before they arise. If there are discrepancies in different databases, the labour inspectors sometimes have to 'coach' employers to support them to use the system effectively and prevent problems. In an e-government context, this type of problem prevention is considered very important.

The starting point to developing this holistic coordinated approach towards databases was set out in Belgium's strategic plan (mission, vision and strategy) which provides the basis for tackling social fraud and undeclared work. This joined-up holistic approach was/is considered an important 'push' factor in setting up a risk assessment system, which is one of the objectives of the plan. To meet this objective, the NSSO initially developed a conceptual methodological framework for risk analysis, which required the identification of risk profiles, the provision of information for prioritisation, supply of information for efficiency gains, as well as data mining capabilities to plan targeted results-oriented inspections. To develop a thorough framework on risk analysis, understanding how each organisation functions has also been necessary, taking into account the specific basic tasks and strategies of the organisation. Business processes were analysed and existing databases and statistics were examined and re-defined where needed. Once general agreement is reached on a datamining model at the conceptual level, this concept needs to be translated into practical tools for inspectors.[2]

1.3 How to detect undetected work

Key messages for policymakers:

A tailored national approach to risk assessment and data-mining is recommended: there is 'no one size fits all' solution, because of differences in resources and ambitions of different countries.

Inspectors need to be adequately trained and have the right skills sets.

A unique identification number for individuals in datasets helps to anonymise data (for data-sharing purposes) and makes the identification of individuals easier and thus facilitates the detection of fraud and error.

The group discussed the Belgian approach to the use of analytics for detection and how it can support inspection work. This can be challenging if the quality of data is poor. Detecting undeclared work requires:

- Extensive business knowledge, which data miners may not possess;
- Complex data integration and application integration;
- Careful feature modelling (indicators, predictor variables);
- Near real-time data; and
- Complex investigation procedures.

One approach to detecting undeclared work could include taking a specific sector, region and timeframe and looking for employers that deliver more value with fewer employees. Comparison can be done with 'peers' in the sector, region, size or even by time (crisis year, boom year), known as

[2] See for example details on Belgium's Miningwatch tool (http://ec.europa.eu/social/BlobServlet?docId=18372&langId=en).

dynamic benchmarking. This can be complemented by more complex investigation procedures (which are legislated for), to interrogate workers about their work at other sites or other times with the same employer. *Under*-declared work can also be detected in this way.

To build predictors, or a predictive model, to assess the risk of undeclared work in employers, data from different applications and institutions is combined. The Belgians have found that it is possible to develop risk assessment approaches without investing in expensive tools, but that it *is* necessary to invest in people and in time. The Cross Industry Standard process (CRISP) can be used in developing a risk analysis system to plan and design, implement, monitor and evaluate a system in a continuous circular process.[3]

There are also challenges in detecting irregularities in declared work – for instance, databases may not detect that one person is the same as another if they have been declared twice by switching the order of their name/surname. Carrying out data quality improvement can help address this and the NSSO uses tools to improve data quality. This also highlights the importance of issuing a unique identification number to each individual. This also anonymises data so it can be easily shared.

1.4 How to use analytics for detection

Key message for policymakers:

Feedback mechanisms can improve analytical tools to create more effective policies and investigations when they are used promptly as well as longer-term.

Effective use of visualisations can help 'join the dots' between various infringements and support better detection.

Belgium presented how they use analytics to prevent undeclared work through the inspection of new employers. Approximately 30 000 'new' employers are registered every year and analytics are used to determine the riskiest employers to inspect. In 2013-14, the sectors targeted were restaurants, construction and cleaning. Six types of risky employer in these sectors were identified. Following an evaluation of its effectiveness, with analytics, the infringement rate of employers inspected reached 52% (without these analytics, it was 40%). The rate is likely to be higher if inspectors implement training on how to register feedback more effectively. Another problem is that it is very difficult to measure the effectiveness of preventive measures. Use of data mining can lead to bigger, more complex investigations, that may yield much higher results in the long term.

Data analytics supports the detection of risks but also:

- Supports investigators by helping to target investigations, case treatment and reporting;
- Enables evaluation or preparation of policy through predictions, ranking and visualisations of connected individuals or transactions (e.g. to spot multiple bankruptcies).

Analytics also support Belgium to control social dumping[4] by using ranked lists of target locations based on 'standard' relative weights for each indicator (which can be adapted). A feedback mechanism enables improvement to the model and reporting to management and government.

Labour inspectors have a crucial role to play: their feedback is crucial at all stages of development. Input from inspectors is needed to define and assess the potential alarms/indicators and to check

[3] Belgium's example from the Catholic University Louvain of the search for letterbox companies in the transport sector on the basis of Orbis, the database of Bureau van Dijk, has demonstrated what is possible as an outcome (http://ec.europa.eu/social/main.jsp?catId=1299&intPageId=4875&langId=en).

[4] Social dumping is a practice where employers typically exploit wage differences between countries or use more vulnerable workers to gain a competitive advantage.

their effectiveness in the field. Involving inspectors at all stages of the data gathering and data matching/mining process is also important to gain their trust and increases the chances that they will embrace the tool. It also takes time and effort to train inspectors for them to integrate predictive tools in their daily work. The Belgian experience suggests that feedback from inspectors is good when it is registered rigorously, with clear deadlines, classified to categories, and time-stamped. They also suggested that three types of feedback are needed:

- Short-term feedback: provided immediately after inspection covering the immediate obvious infringements;
- Mid-term feedback: after the conclusion of the inspection 'case';
- Long-term feedback: including the outcome of legal cases.

Belgium supports inspection work with analytics involving geo-visualisation and network visualisation tools[5]. Geo-visualisation (through Google Maps) enables undeclared work to be tackled in real-time, for example, by examining the relationships and connections between fraudulent companies, premeditated bankruptcies, etc. Belgium has also developed other examples using geo-visualisation and network-visualisation.

1.5 Evaluation

Key message for policymakers:

Return on investment (ROI) should be clearly measured to help enforcement bodies understand the impact of more advanced data usage on the outcome of inspections

Developing balanced scorecards on the basis of key performance indicators to measure the output of predictive models can increase the effectiveness and impact (as well as ROI) of an inspection plan based on datamining and risk analysis

An evaluation of the return on investment (ROI) of these systems shows that while the cost of investing in this system is high, the benefits in terms of the impact on preventing and reducing undeclared work is also high, including other types of social fraud e.g. social dumping.

Belgium has found that the ROI should be clearly measured in order to help decision-makers understand the positive impacts of more advanced data usage on the outcome of inspections. Such ROI evaluations increase internal and external awareness of the capacities of public administrations and the potential value of investments in data sharing/matching/mining in order to prevent and deter non-compliance with labour and tax rules.

To increase the effectiveness and impact (as well as ROI) of an inspection plan based on datamining and risk analysis, methods for measuring and evaluating results should be developed, such as balanced scorecards on the basis of key performance indicators. To measure the output of predictive models, benchmarks can be agreed amongst partners, inspectorates and institutions. One criterion to assess the effectiveness could be to measure the rate of transformation from undeclared work to declared work in a given company, sector or timeframe, and the economic and social benefits of doing this.

1.6 Conclusions

Belgium has demonstrated that support from key influencers is needed at all levels of government. To make data analytics effective, it also requires highly skilled people, knowledge of the businesses /

[5] See for example a practice examples on Checkin@work on high risk construction companies http://ec.europa.eu/social/BlobSe0rvlet?docId=18322&langId=en

sectors involved, and a supportive culture which understands that risk analysis can make the core business of an inspectorate more efficient. This process takes time, and an effective feedback procedure is also needed.

Other lessons from the Belgian experience include:

- The first step in developing a system for risk analysis can be a pilot project (i.e. a small and manageable project), gradually moving forward from that, step-by-step.
- The key questions triggering the development of a risk analysis system should be: a) what is the fraud phenomenon that needs to be tackled? and b) which instruments, statistics and databases are already available as a basis for identifying and extracting red flags and indicators that can lead to detection of this phenomenon?
- It is important to share and learn from others' experiences and consider the data analysis tools used by other enforcement bodies at national or cross-national level.
- Incomplete and incorrect data records should be reported to the data source holder, who should then take action to correct it and reduce 'dirty data'.
- Defining risk indicators and red flags is an important but difficult exercise, so adequate time and resources need to be devoted to this.
- If a predictive model is to be developed, it should be fully integrated into the business and functioning of the enforcement authority (top-down/ bottom-up, including continuous assessment and evaluation of its effectiveness). Belgium's experience demonstrates that results improve over time and through continuous evaluation and re-adjustment.