

Modernising Labour Market Statistics: Administrative Data Integration and Survey Redesign

Katja Olofsson, Statistics Sweden, katja.olofsson@scb.se

Susanne Gullberg Brannstrom, Statistics Sweden,
susanne.gullbergbrannstrom@scb.se

Abstract

This document presents two complementary projects undertaken by Statistics Sweden to modernise labour market statistics, focusing on both fulfilled labour demand and recruitment needs. The projects share a common aim of improving statistical quality and relevance while reducing response burden and production costs through the combined use of administrative data and redesigned surveys.

Part I describes the integration of administrative data into the production of employment statistics. By consolidating three previous statistical products into a single system based on individual-level employer declarations from the Swedish Tax Agency, the project replaces survey-based data collection with administrative sources wherever possible. The new appr improves consistency over time, enables monthly statistics at a detailed regional level, and aligns more closely with national accounts. It also incorporates additional data, such as sick-pay information, while significantly lowering respondent burden and production costs.

Part II presents the redesign of employer surveys on labour demand, resulting in the new survey Job Openings and Recruitment Needs. The project emphasises key methodological principles, including the distinction between interest, target, and observation variables, as well as the importance of iterative development informed by both user needs and respondents' practical reporting conditions. The redesign led to a substantial reduction in response burden—nearly 40 per cent—while maintaining and extending the relevance of the statistics.

Together, the projects demonstrate how integrating administrative data with carefully designed surveys can produce more efficient, robust, and user-oriented statistical systems. They highlight the importance of methodological clarity, user engagement, and organisational collaboration in the ongoing modernisation of official labour market statistics.

Keywords: labour market statistics, respondent burden, user needs, survey redesign, data integration

Introduction

This document presents the outcomes of two closely related projects aimed at modernising labour market statistics in Sweden. Both projects were developed through a joint effort at Statistics Sweden, with the shared objectives of improving statistical quality, increasing efficiency, and reducing both respondent burden and production costs. While the projects focus on different aspects of labour demand, they are unified by a common methodological perspective and a strong emphasis on user needs, data availability, and sustainable statistical production.

Part I focuses on the measurement of fulfilled labour demand through the integration of administrative data. By replacing and consolidating several existing statistical products, the project establishes a new production system based primarily on individual-level data from administrative sources. This transition enables more consistent and timely statistics, while reducing the need for direct data collection.

Part II complements this work by addressing the measurement of labour demand through survey redesign. It describes the development of a unified employer survey, highlighting key methodological considerations such as the distinction between different types of variables, the importance of iterative development processes, and the need to balance user requirements with respondents' reporting capabilities.

Together, the two projects illustrate a broader transformation in official statistics, where administrative data and targeted surveys are combined to create more coherent, efficient, and user-oriented statistical systems. Their joint results contribute to a more comprehensive and modern framework for labour market statistics in Sweden.

Modernising Labour Market Statistics

Part I: Administrative Data Integration – Employments

1. Introduction

Labour market statistics in Sweden have evolved over time, resulting in several statistical products that are not fully comparable due to differences in design, data

collection, and output. At the same time, increased access to high-quality administrative data has created new opportunities to modernise production.

Against this background, Statistics Sweden initiated a project to integrate administrative data into the estimation of fulfilled labour demand. The project aimed to replace and consolidate three existing statistical products into a single, more efficient system covering the number of employees, labour costs, hours worked, and sick pay during the employer-paid period. A key constraint was to implement these changes while maintaining EU reporting obligations and avoiding an extended period of parallel data collection.

2. Aim

The main objectives were to reduce respondent burden and lower production costs, while improving the accuracy and consistency of statistics. This was achieved by replacing survey-based data with administrative data, particularly the monthly PAYE tax returns at individual level collected by the Swedish Tax Agency. In addition, sick-pay data, collected through a sample survey, were integrated into the same framework.

3. Method

The project followed a structured and collaborative approach. A transition plan was developed in close dialogue with Eurostat to ensure compliance with Short-Term Statistics (STS) requirements, particularly regarding the definition of *“Number of persons employed”*.

A comprehensive user consultation process was conducted, including targeted seminars and workshops with key stakeholders such as government ministries and central economic institutions. Two main messages guided the dialogue: to inform that not all previous variables could be reproduced with the same definitions, and to find out what information the users needed.

Based on these insights, an administrative dataset was constructed using PAYE tax returns, including information on wages, benefits, taxes, and employer contributions. These data were supplemented with business information.

Two key modelling components were developed:

- Number of hours worked were estimated by combining monthly wage statistics with administrative data on employees.
- Regional allocation of the estimated values was performed using a top-down approach based on the distribution of local units and the number of employees according to the Statistical Business Register.

Finally, as no comprehensive administrative source on such data exists, the survey regarding sick-pay data, was redesigned to better reflect user needs and to better align to the register-based number of employees.

4. Results

The new system has resulted in several enhancements:

- Improved accuracy and consistency, with more stable time series
- Monthly statistics with regional detail, replacing quarterly outputs
- Better alignment with national accounts, resulting in improved coherence
- Reduced respondent burden, as one survey was fully replaced
- Lower production costs

The new statistics estimated higher employment levels in certain industries, particularly such as restaurants, suggesting that earlier statistics may have underestimated employment since the survey design didn't fully capture staff on weekends. On the other hand, the new approach might overestimate employment, since everybody who has received salary during the month is counted for.

5. Discussion

While the results are considered successful, further improvements are possible. The current top-down regional allocation method does not capture differences between workplaces, such as gender or age distributions. Future work will explore the feasibility of a bottom-up approach using workplace-level information reported by employers.

Additionally, reducing the reliance on model-based estimation of hours worked remains a priority, with the long-term aim of deriving these variables directly from administrative sources.

Part II: Survey Redesign – Labour Demand

For several years, three separate surveys on labour demand were targeted at employers in Sweden, two in full and one in part, by Statistics Sweden. Individually, each survey was well justified, but in combination they resulted in undue response burden and suboptimal production.

Against this background, Statistics Sweden and Arbetsförmedlingen (the Swedish Public Employment Service) carried out a project in 2022–2024 to design one unified employer survey within the area of labour demand statistics. The aim was to a) reduce response burden, b) meet new user needs for statistics on labour and skills demand, and c) increase efficiency in statistical production. The project was carried out by a multidisciplinary team consisting of six subject specialists (one acting as project leader), four methodologists (including three statisticians and one cognitive specialist), and representatives from the data collection function. Five of these worked full-time or near full-time for most of the project period.

The development process included two preliminary studies – one on respondents' reporting practices and one on users' statistical needs – along with extensive remodelling of the sample design, a two-step evaluation of the questionnaire, and a complete redesign of the data collection process. The result was the new survey Job Openings and Recruitment Needs (LOR), launched in 2024, which reduced respondent burden by nearly 40 per cent, while also meeting both existing and newly emerging statistical needs.

1. Distinguish between interest variables, target variables and observation variables

A rather common approach among analysts, both within and outside National Statistical Offices (NSO), is to begin survey design by outlining questionnaire items. Instead of first formulating the target variables, attention is directed towards the wording or scope of individual survey questions. This tendency is particularly evident when collaborating with non-NSO organisations. In such situations, representatives of the NSO have a key responsibility to clarify the distinction between interest variables,

target variables and observation variables, and to redirect the discussion towards the target entity and its specification.

An example from the LOR project is the concept and measurement of *immediate demand for labour*. During the user consultations, several government agencies expressed a need for statistics on a variable with a shorter time frame than *job openings*, as the latter may, for example, also capture seasonal demand for labour at a later point in time. This need was described in terms of “acute demand” or as a variable reflecting the “real need for labour here and now”. In other words, this represented the **interest variable** in this part of the statistical domain and, to a large extent, corresponded to the existing **target variable** *vacancies* (distinct from job openings), defined as “unfilled jobs for which staffing is desired on the reference date”. This target variable was therefore retained unchanged in the new survey.

However, the preparatory study with respondents showed that the earlier question, which directly reflected the formal definition of the target variable, was perceived as too abstract and difficult to answer. Therefore, the observation variable was reformulated as the question: “Approximately, how many of these [number] job openings required staff to start immediately or as soon as possible?” The observation variable was discussed and validated with core users as a sufficiently accurate proxy for the target variable.

2. Adopt an iterative approach when evaluating user needs and respondents’ actual prerequisites

The traditional process in statistical production and in the development of new surveys is often to first establish user needs and then assess data availability. This is, for example, the suggested approach according to Statistics Sweden’s support for the statistical production process. However, if this workflow is followed strictly, there is a risk that discussions with statistical users do not take the conditions for data provision into account. For instance, the risk increases that user consultations engage in detailed discussions of variables that, in practice, will never be feasible to collect. An example from the LOR project is the target variable “number of agency workers”, which was included in one of the earlier surveys. The preparatory study with respondents clearly

showed that especially larger establishments were unable to answer questions on this theme reliably. In many cases, aggregated invoices covering total costs for agency staff were the only available source. The variable was therefore excluded from the new survey.

Overall, the project benefited considerably from starting with a preparatory study involving respondents, as it served as a realistic backdrop for all subsequent development.

3. Reassess the balance between user needs and respondent burden in each design decision

The requirement to balance user needs against response burden is central to any survey and is also explicitly stated in the Swedish Official Statistics Act (SFS 2001:99). The principle is to be applied at the aggregate level, in terms of overall user value and total response burden. However, in practice, striking a balance is equally important for smaller design decisions, such as question wording, and the handling of respondent cases in the survey.

An example of a larger design decision in LOR is that data collection for the variable vacancies, reflecting a trade-off between need for data and response burden, was changed from monthly to quarterly for all units. An example of a smaller design decision is that establishments unable to report for their entire operations are allowed to report for a clearly dominant part, even if this results in some systematic measurement error.

4. Acknowledge that some development tasks may lack clear organisational ownership

Designing new, large-scale surveys with mandatory response requirements is not a frequent task at Statistics Sweden. It requires a project team with diverse competencies, where all members assume both practical and psychological responsibility, i.e. a strong sense of ownership. Even within the operational management of a survey, it is, regardless of organisational structure, inevitable that some tasks or issues do not clearly fall within a single role. In practice, responsibility may be diffuse, with some tasks not clearly assigned and others overlapping across

roles. In a larger development project, such challenges become even more pronounced.

A key part of the solution is extensive and transparent communication, collective responsibility, and an understanding that the survey is carried out by the NSO as a whole, rather than by any single part of the organisation. Another important factor for effective collaboration is leadership that encourages staff autonomy, while treating challenges as shared concerns of the NSO rather than as issues confined to individual organisational units.

Overall Conclusion

Together, Part I and Part II demonstrate a coherent strategy for modernising labour market statistics through the combined use of administrative data and carefully redesigned surveys. The projects highlight the importance of methodological clarity, user engagement, and organisational collaboration in building efficient and sustainable statistical systems.

List of references

Statistics Sweden (n.d.) *Employments*. Available at: <https://www.scb.se/en/finding-statistics/statistics-by-subject-area/labour-market/labour-force-demand/employments/>

Statistics Sweden (n.d.) *Job openings and recruitment needs*. Available at: <https://www.scb.se/en/finding-statistics/statistics-by-subject-area/labour-market/labour-force-demand/job-openings-and-recruitment-needs/>

Statistics Sweden (n.d.) *Development of labour market statistics*. Available at: <https://www.scb.se/dokumentation/statistikomraden-under-forandring/oversyn-av-arbetsmarknadsstatistik/>

Statistics Sweden (2021) *Subject Area Design Labour Market – Final report*. Available at: <https://www.scb.se/contentassets/4151635ea4b74b22a3ba8ecc0c81874f/subject-area-design-labour-market-final-report-may-2021.pdf>