

An Eleventh Application of ASPIRE for Statistics Sweden

Susan Linacre, Stephen Penneck, Laurie Reedman

September 29, 2022

Contents

1. Executive Summary	3
2. Introduction	4
2.1 The current review round	4
2.2. Reflections	4
3. General observations and cross cutting recommendations.....	5
3.1 General Observations.....	5
3.2 Cross Cutting Issues and Recommendations	5
3.2.1. Statistical leadership and user focus	5
3.2.2. Methodologists	6
3.2.3. Analysis.....	7
3.2.4. Managing change	7
4. Product scores and recommendations	8
4.1. Interpretation of the scores	8
4.2. Production Value Index (PVI)	10
4.3. Quarterly Emissions Accounts (QEA)	13
4.4. Building Statistics	16
4.5. Population statistics	19
4. References	22

1. Executive Summary

Three previously reviewed products were reviewed in 2022: Production Value Index (PVI), Quarterly Emissions Accounts (QEA) and Building Statistics, as well as one new product, Population Statistics. Section 2 gives the background to ASPIRE. There were no changes this year from the previous round. The key findings are presented in section 3 and cross-cutting recommendations are presented in section 4.

All product areas showed a high level of professionalism in relation to their products and were constructive and supportive throughout the review process. We were all grateful to hold interviews face to face at the Solna office of Statistics Sweden this year.

The Production Value Index is moving towards less reliance on survey data and increased use of tax data, particularly VAT data. Work has been done to explore the quality of VAT data and to compare results using the previous and new methods. We encourage the team to foster a strong relationship with the tax authorities, to better understand the quality of tax data and to bring about mutually beneficial quality improvements.

The Quarterly Emissions Accounts are derived from models applied to survey and non-survey data. The production team have built quality assurance mechanisms into their production process and are constantly striving to improve the precision and effectiveness of model parameters. While they have a good sense of the quality of the statistics across industries, there is room for improvement in measuring and reporting the accuracy and uncertainty in their statistics.

The main quality concern for the Building Statistics program continues to be the lag in capture of data from municipalities. Staffing shortages have exacerbated this problem and redevelopment of the front end of the processing system offers the opportunity to gain substantial improvements in both efficiency and quality. This is likely to require corporate input both in terms of resources and to build sufficient influence with municipalities. There is also an opportunity for Statistics Sweden to play a statistical leadership role and build their own understanding of the value of the monthly seasonally adjusted permits data as a leading indicator of economic activity, and if it proves valuable, to publish it for general use to inform discussion about current economic conditions, as is done in several other countries.

The Population Statistics program was reviewed for the first time this round. The production team maintains an ongoing relationship with users through the User Board, whose members appear generally content with the quality of the statistics. However, it is of concern to the ASPIRE team that, while the richness of the Total Population Register lies in its ability to provide data at the regional level for various population subgroups, there is very little evaluation of the data below the total population level.

Cross-cutting issues that surfaced this year related to the need to:

1. Continue to build statistical leadership skills and expectations across the organisation.
2. Develop richer relationships with users, becoming intelligent suppliers with an understanding of the key questions users need answered with effective statistical solutions, and supporting them in becoming intelligent customers.
3. Build the methodological role in the organisation to support increased analysis, and the development of new methods and techniques within business areas, and to help in the identification of opportunities for, and uses of, corporate tools and processes across the organisation.
4. Build analytic skills in business areas to allow the exploration and understanding of their data, and its quality, for example in undertaking revisions analysis or deciding how best to present the information inherent in the statistics in a meaningful way to users.
5. Include in the change management plan for the new editing strategy guidance on decision making and contingency planning for programs in lesser states of readiness to absorb the transition in their regular budget and resource level.

2. Introduction

In 2011, the Ministry of Finance directed Statistics Sweden to develop a system of quality indicators for a number of key statistical products. ASPIRE was developed to meet this need, and conducted annually from 2011 to 2017.

A review of the ASPIRE process was done in 2018 and some changes were implemented in 2019-2020 which affected the scoring from 2020 and onwards. Five of the ten previous products were rotated out in order to provide space for new products which were first evaluated in 2020. A two year period between review rounds was created in order for the products to have enough time to work on recommendations.

The 2020 and 2021 review rounds differed to previous rounds with remote assessment due to the COVID-19 pandemic.

2.1 The current review round

The current round was undertaken with a physical visit to Statistics Sweden in Solna. This allowed more in-depth discussions with both the product areas and top management.

The reviewed products were the same as those reviewed in 2020 i.e. Production Value Index (PVI), Quarterly Emissions Accounts (QEA), Building statistics. An additional product, Population statistics, was added for the current round.

The product teams did their necessary preparatory work with the checklists and quality documentation which was sent to the expert team in advance of the review as usual. Thereafter, the review team had a written exchange with follow-up questions and answers with each of the product teams before the review.

The external review team was made up of three experts – Susan Linacre and Stephen Penneck who had joined the team in 2018 and Laurie Reedman who joined in 2021.

The review team would like to express their gratitude to staff in Statistics Sweden who have participated in ASPIRE round 11. We are also grateful for the quick responses we got to our questions and for the open way that staff responded.

2.2. Reflections

The ASPIRE team queried whether Building Statistics as defined in the current scope warrants full inclusion in ASPIRE every two years. While important, the statistics covered are not a significant component of Statistic Sweden's work program, and the area has limited access to resources with which to drive change. Instead of full inclusion it is suggested that the frequency for a full review of Building Statistics be dropped to 4 yearly, with a review of progress against recommendations made every two years.

The ASPIRE team also questioned the scope of the Quarterly Emissions Accounts review. In future perhaps it would be beneficial to include Annual Air Emissions as well, since the two programs are closely related.

In terms of the review of Population Statistics, it was noted that a number of areas in Statistics Sweden (such as the Microdata Service Group) are involved in the production of statistics from the Total Population Register (TPR). Much of the documentation on quality provided to the ASPIRE team related to population estimates at the national level. To get a better view of the quality of population estimates at a subnational and subpopulation level, consideration might be given to broadening the scope of the review to include analytic outputs from the TPR.

The ASPIRE team noted that among the three programs reviewed this time that had been reviewed previously, only roughly a quarter of the recommendations had been completed. This raised the question of whether the programs under review are the most relevant to Statistics Sweden given their current budgetary constraints, or if budget decisions take into account the resource requirements for addressing ASPIRE recommendations.

3. General observations and cross cutting recommendations

3.1 General Observations

We learned during our review about Statistics Sweden's new strategy, with its five overarching goals and specifically the Operational Plan for 2022-24. The goals seem relevant to many of the priorities of the ASPIRE programme, especially the focus on the useability of statistics and the need for the agency to pursue staff development, and will provide an aspirational tool to advancing the statistical leadership of the office. We heard that the new strategy will enable Statistics Sweden to better set priorities and noted the new priorities of labour market statistics, economic statistics, communication and secure data supply.

We noted that the new goals are essentially also goals for the Swedish Official Statistical System, which is made up of all the statistics agencies in Sweden, and form an excellent basis for the forthcoming discussions between the DG and his counterparts.

The new strategy is underpinned by the new organisational structure introduced last year, designed to provide greater clarity of responsibilities and to put methodology more at the centre of the organisation under the new Deputy Director General. Given this is being done at a time when working from home is being formally recognised for the future, there will be a need to support the new Section heads - effectively the change leaders - in their new roles.

The recent Eurostat Peer Review report of the Swedish Statistical System presented a number of challenging recommendations, on some of which Statistics Sweden has expressed diverging views, including three relevant to some of the issues we have raised in the past (R1 on statistical leadership; R2 on priority setting; and R20 on user consultation). We can see that Statistics Sweden is constrained by the institutional framework under which they operate and were encouraged to note that although they were not able to accept the recommendations in the terms in which they were set out, nevertheless the office will find other ways of addressing the underlying concerns raised by the report.

We were encouraged to hear of some of the new initiatives that will enhance the agency's capability. The new data centre will improve data security, and the new data platform – if financially supported – will provide additional analytical capability for users. It will be important to build stakeholder support for the benefits this will bring.

Most countries have experienced non-response issues with their household budget surveys, which have impacted on quality of their principal uses – as major components of the national accounts and as weights for the consumer price index. Since 2017, ASPIRE has been recommending steps should be taken to look at replacement sources, possibly from the supply side, including scanner data. We were pleased to note the progress being made. The new approach uses scanner data, credit card transaction data and other sources to estimate household consumption. The ASPIRE team noted the challenges that there will be in bringing these data together into a consistent framework, and in making estimates for missing areas. The project team are confident of delivering estimates for national accounts use by 2023 and we look forward to reviewing progress next year as part of the GDP/ CPI product evaluations.

Statistics Sweden took the opportunities presented by COVID to discuss the trade-off between accuracy and timeliness with users, and introduced innovations, including a quarterly flash estimate of GDP, and an analysis task force. The office will want to continue to exploit such approaches in the future.

Next year sees Sweden take on the Presidency of the European Council which will add to the pressures on the agency, but also provide them with opportunities to shape the European statistical agenda and showcase their achievements.

3.2 Cross Cutting Issues and Recommendations

3.2.1. Statistical leadership and user focus

Statistical leadership has been a theme of our last two reports and we are pleased to see that the recommendations in last year's report have been taken forward. As a result the idea is beginning to gain currency in the office. It is referenced in the Statistics Sweden strategy and forward plan. Importantly it is now a term more widely recognised and used in the office at working level.

But there is as yet less evidence of statistical leadership being widely displayed at product level, although of course there are exceptions. Especially in relations with users we are looking for Statistics Sweden to understand the business of their customers – knowing what the policy questions are that people want answers to – so that they can think about what best meets their needs, and to what quality standard. In this way Statistics Sweden becomes an ‘intelligent supplier’, and by having discussions about needs and quality standards with users, users can similarly become ‘intelligent customers’.

We noted that Statistics Sweden is about to move its customer focus towards the expert user, and this is an opportunity to think through how best to establish this richer relationship.

But statistical leadership is not just about the relationship with users. We would expect statistical leaders in product areas to overcome organisational barriers (‘silos’), taking overall responsibility for the quality of statistics in their area and to work in wider teams to take forward this agenda. The leadership role of the methodologists fits in here.

We encourage Statistics Sweden to:

1. Develop a programme of workshops with product areas and methodologists so that the statistical leadership expectations are clear to them, and to include the concept of statistical leadership in any management and leadership training offered.
2. Use the opportunity of moving the customer focus towards the expert user to develop a richer relationship with them which enables Statistics Sweden to become an intelligent supplier, and the expert user to become intelligent customers.

3.2.2. Methodologists

We saw that the role of Methodology in driving the strategic directions of Statistics Sweden over the coming years is recognised in the Operational plan and the new organisational structure to support it. We welcomed this as the methodologists will help the organisation in its move to an effective and efficient corporate approach using common tools and methods. They will act as a force for integration across the various stove pipes, identifying opportunities for common approaches and helping develop the methods and tools needed. They will also be able to help build the analytic and data exploration culture and skills of the organisation.

As the Methodology group moves from a bottom up service provision role to a more strategic approach, it will be important for their role, including expectations of them (i.e. their accountability), and their authority to be clarified. The size of the methodological group is not large, and in a tight labour market, recruitment and retention will be difficult. It will be important to invest in developing and nurturing the group as a whole, for example through cross organisational networking, staff rotation, and seminars. The statistical leadership capability of methodologists should also be developed through example, support, and leadership training. The overall work program of the methodologists in the organisation should be transparent and driven with a ‘top down’ perspective, in partnership between the head of Methodology and Architecture and the Economic and Social Statistics Departments. This will provide corporate assurance that strategic directions are being appropriately supported, and methodological resources are set at an appropriate level.

We encourage Statistics Sweden to:

1. Develop an explicit strategy for building the methodological capability as an organisational resource, in particular paying particular attention to a common recruitment, training and development strategy across the organisation, to build both the technical and statistical leadership skills of the group.
2. Develop the methodology work programme as a top down and transparent process, aimed at driving strategic change as well as meeting business continuity needs.

3.2.3. Analysis

Improved analysis is one of the strategic directions of the Statistics Sweden Operational Plan. This analysis typically covers developing new analytic products such as the new Labour Market Statistics, and analytic outputs answering specific key user questions. Equally important is the more internally focused day to day exploration of data sources and statistics, to understand them, understand their quality, and to present them in a way that is most meaningful to users.

Building the analytic capability of the organisation will take time, and involve in part the appropriate recruitment of skills. However an important part of the strategy will be to build confidence and skills in accessing and exploring data at the day to day business level, supported by methodological resources. This may be a question of providing some space in the work program for such exploration, encouragement to be curious about the data, and sharing of skills and experience across subject areas, even for small examples of day to day analysis. Impediments to such exploration of the data, such as ease of access to the data or to analytic tools should be identified.

We encourage Statistics Sweden in its intention to increase analysis within the organisation and in particular note the need to develop the confidence and skills in the use of analysis in day to day business of statistical areas by:

1. Providing encouragement and space in the work program to undertake relevant work, such as quality exploration, revisions analysis, and analysis relating to how best to present data to highlight patterns within it.
2. Identifying and addressing any impediments to such analysis relating to data access and analytic tools, as well as skills.
3. Sharing analytic experiences through seminars and networks.

3.2.4. Managing change

Statistics Sweden is phasing out manual editing of economic survey data in order to use resources more efficiently while rendering data more fit for purpose. Although this initiative is driven by financial considerations, it is also an opportunity to better understand and to improve data quality. It is important for Statistics Sweden to clearly identify specific quality related objectives for this modernization activity, otherwise they risk an outcome in which editing is done differently but their desired goals are not achieved.

Through the ASPIRE interviews it was noticed that programs vary in their state of readiness to transition to the new editing strategy. While some programs are already using common tools and anticipate minimal disruption, other programs have outdated and fragile processing systems, where the risk of error or even breakdown is significant. The change management plan needs to have business continuity and risk mitigation considerations for programs where implementation could involve redevelopment of more than just the editing modules. Another new strategy that will be explored more in future ASPIRE rounds will be the creation and use of a central frame for economic surveys. This is another example of a modernization that will succeed through careful change management.

We encourage Statistics Sweden to:

1. Include in the change management plan for editing guidance on decision making and contingency planning for programs in lesser states of readiness to absorb the transition in their regular budget and resource level.

4 Product scores and recommendations

4.1. Interpretation of the scores

There is a natural tendency to compare the overall scores across the products or to rank the products by their total score. However, the ASPIRE model was not developed to facilitate such inter-product comparisons and there are some risks associated with ranking products in this manner. For one, the average score for the component, Sources of uncertainty, for a product reflects a weighting of each single source of uncertainty by their importance to Overall accuracy, which can vary considerably across products. Products with many highly important sources of uncertainty may be at somewhat of a disadvantage in such comparisons because they must perform well in many important areas in order to achieve a high score.

Furthermore, the assessment of low, medium, or high importance to overall accuracy is done within a product, not across products. Thus, it is possible that a highly important source of uncertainty for one product could be of less importance to Statistics Sweden than a medium important source of uncertainty for another product if the latter product carries greater importance to Statistics Sweden or for official statistics. If resources devoted to accuracy improvements are greater for one product than another, this could also explain why some products are able to show greater improvements than others. Further, although we have attempted to achieve consistency in ratings among products, some inconsistencies surely remain.

Finally, the scores assigned to a particular source of uncertainty for a product have an unknown level of uncertainty due to some element of subjectivity in the assignment of ratings as well as other imperfections in the rating process. A difference of 2 or 3 points in the overall product scores may not be meaningful because a reassessment of the product by different reviewers could reasonably produce an overall score that differs from the assigned score by that margin. Thus, any ranking of products would need to acknowledge these inevitable and unknown uncertainties in the ratings.

Normally, a more appropriate use of the product scores is to compare scores for the same product across review rounds as a way of assessing progress toward improvements. In the case of the reviewed products in round 11, this would involve comparing scores with those for round 9 when they were reviewed last.

Table 1 shows the summary scores for the four products that were reviewed in round 10.

Table 1. Summary of Average Scores by product

Product	Overall accuracy (average scores)	Sources of uncertainty (weighted average scores)
PVI	40	40
QEA	47	47
Building statistics	35	37
Population statistics	43	43

Table 2 shows the average scores per product for each component of Accuracy. The importance of the single sources of uncertainty to Overall accuracy - high, medium, low or not applicable - is indicated by the shaded cells. The average scores for each Accuracy component across the three products are shown in the second last column together with the weighted average scores in the last column. The weights of 3, 2, 1, and 0 correspond to the categorisation of high, medium, low or not applicable regarding the importance to Overall accuracy.

Table 2. Average scores by Accuracy component and product

Sub and sub-subcomponents of Accuracy	PVI	QEA	Building	Population	Average score	Weighted average score
Overall Accuracy	40	47	35	43	41	N/A
Sources of uncertainty:	40	47	37	43	42	N/A
-Sampling	32	40	N/A	N/A	36	35
-Frame coverage	25	53	35	42	39	39
-Measurement	52	42	38	40	43	43
-Non-response	42	42	35	47	42	42
-Data processing	45	50	40	48	46	45
-Model assumptions	47	48	38	43	44	45
Preliminary statistics compared to final statistics	40	33	35	40	37	N/A

Importance to Overall accuracy			
N/A	Low (L)	Medium (M)	High (H)
Weights			
0	1	2	3

4.2. Production Value Index (PVI)

Context and Discussion

The PVI is a monthly index showing the economic progress of the business sector, broken down by industrial sector. The quarterly index is inventory adjusted, and is an input into GDP and the national accounts. Surveys currently provide the main sources, though survey data is replaced by tax data for smaller businesses as it becomes available, and so the survey data is especially relevant for the early estimates and for the estimates for larger businesses. Overall accuracy is about fitness for purpose, and the main purpose of the PVI is as an input to the national accounts. There is good dialogue between the PVI product area and the national accounts, whose interest is in quarterly growth rates. Other users are the Riksbank and the National Institute of Economic Research (NIER), which are again interested in growth and the identification and assessment of turning points in the economy.

At the last (2020) review we expressed a concern about the lack of information about how these institutions use the statistics or what their quality requirements are. Our concern was that the current approach of focusing on year-to-year movements, while providing a crude seasonal adjustment and trending mechanism, is suboptimal in identifying movement at the current end of the series. We recommended that the product area maintain a more active dialogue with the main non-national accounts users to ensure they are making the best use of the statistics and that the product area understands the uses to which the statistics are being put.

We were pleased to note in the current review that a meeting had been held with the NIER in 2020, who had confirmed they use the statistics for forecasting, but this had not illuminated the area of concern raised in 2020, nor had any more recent meetings been held.

However, work is underway to improve the existing publication, led by the Communications department. We hope this will enable a focus on graphical representation of time series, including seasonally adjusted series, and month to month movements.

We had also recommended developing measures relevant to the accuracy of movements in the series. The product area told us that it would be possible to produce measures for the sample and the frame, and they had studied the literature on identifying turning points, but this work had not been progressed due to lack of resources.

The PVI is highly dependent on model assumptions. Other sources of uncertainty in the estimates (sampling, frame coverage and measurement) make only a moderate contribution to overall accuracy, with a low contribution from non-response and data processing. Sampling has an impact on the early months of the estimates, before the survey is replaced with tax data for smaller businesses. These early months are important in economic assessments as it is when data are being scrutinised for turning points in the economy. The sample is selected from the new register at the beginning of the year and then is maintained across the year. We comment further on this below.

The frame has a cut-off to exclude the very smallest businesses. This is standard practice, from a pragmatic perspective, but it is important to evaluate the level of the cut-off in terms of contribution to movement, particularly at times when the level of economic activity is changing, ideally across a business cycle. We were pleased to see that work to evaluate the impact of the cut-off had begun, but we were disappointed that it was discontinued due to lack of resources.

In our last review we had urged a more systematic approach in some problematic industries, such as construction, with any relevant conceptual issues being addressed. We were pleased to see that a new model has been adopted for construction, based on projections of VAT data. We were told that some differences with national accounts estimates remain so there is still further work to be done. New models based on turnover are being used for some mining and transport industries. We were pleased to see that a comparison with existing results had been done before implementation. These developments will contribute to improved accuracy for the PVI.

During the current review we noted other improvements now in place, including an improved process for producing deflators and re-parameterisation of macro-editing.

We were also pleased to see that, following our 2020 recommendation, work has been carried out to look at the differences between survey data and tax data, and this has shown a close correspondence. Following on from this, we were interested to hear of the progress being made to use VAT data instead of survey data. The new source is being run in parallel with the existing results for businesses below the 1:1 stratum this year, and is showing good results. It is expected to be implemented next year, when work will begin to investigate some of the service industries. We look forward to hearing more about progress on this important project at our next review. The development should reduce uncertainty due to sampling on overall accuracy, but the product will become more dependent on measurement error from the tax data, and the team will need to be aware of the lags in receiving these data.

We heard that there has been no discussion with the tax authorities on the quality of VAT data, due to confidentiality concerns. Nevertheless, the product area recognises that there are quality issues both regarding the industry classification, and in the turnover data itself – which does not meet national accounts definitions. For larger businesses this can be corrected by profiling, but not for the medium size businesses. This means that more detailed aggregations may be more subject to measurement error and Statistics Sweden will need to consider carefully the level at which the new statistics can be published.

The product area will want to devise a mechanism for monitoring measurement error in the tax data. This can be expensive to do, and it may be that they will only be able to set up studies every three years or so.

The plan is to continue to sample from the frame for businesses above the small business cut-off but below the 1:1 large business stratum, but then use VAT turnover data as the data source instead of survey data. Of course this means that the practice we commented on at the last review, of only updating the frame annually, remains.

As we said in our 2020 report, a better practice would be to pick the sample quarterly or even monthly if a frame were available. The annual selection of the frame means that businesses born during the year cannot be represented in the sample, nor can businesses that grow above the cut-off. This affects the ability to measure rapid changes in growth rates, through the year. This might be a growing issue, given the greater interest in short term movements in economies, following the recovery from the Corona pandemic and possible impacts of the war in Ukraine and the energy crisis. If new businesses are a driver of growth, or if smaller businesses grow proportionately more quickly than larger ones, then this effect might be missed.

An aspect of the use of tax data for the entire ‘sampled’ portion of the population relates to how to deal with defunct units, and units for which tax data is not yet available. It was not clear what evaluation is underway in relation to this. Known defunct units in the sample are easily handled as nils, and represent other defunct units in the population. It may be that businesses for which tax data are not available, are defunct, or it may be that the businesses are late in reporting the data. The current evaluation should include evaluation of how effectively the proposed process is able to differentiate between these two scenarios, and the impact of any error associated with the difference impacts estimates of movement.

The PVI is in the first phase of the corporate editing project. PVI uses a common software system, which is well controlled, and it will be relatively straightforward to implement the new approaches. The product already uses selective editing and macro-editing. The new approach will enable decimal errors to be more easily corrected, and there will also be benefits from respondent editing. The product area will need to consider what the editing approach will be for the tax data. We were pleased to hear that separately from this strategic project, the product area will continue to make tactical improvements to existing editing approaches.

Revisions analysis is one of the most common tools used by users to assess the quality of preliminary statistics. The product area evaluates revisions as part of its quality control procedures but we were concerned in 2020 that there was no regular publication of revisions time series. We were pleased in the current round to be able to review a new revisions paper, which we understand is to be updated annually. We suggest that a graphical time series presentation of the results would make it easier to understand what is happening. There appears to be a pattern of systematic long run negative revisions emerging. Such patterns need to be evaluated, as they may point to opportunities to improve estimates, for example through improved model assumptions, and they will enable users to make quality assessments. It would also be helpful to extend the analysis back further in time, if possible.

We noted some good practice in holding regular workshops, on seasonal adjustment, prices, editing and quality, which will have increased the skills and capability of the team, but this has been offset by the loss of two key members of staff. The area benefits from an almost full time methodologist, with the work programme agreed at Section Head level. Last year's reorganisation has taken some time to bed in.

Recommendations for the coming two years

1. Continue to engage actively with all users to better understand uses and which measures would best meet their needs; and in the work taking place to improve the PVI publication, focus on a time series graphical representation of the statistics, including seasonally adjusted series and month-to-month movements.
2. As the PVI becomes more dependent on VAT as a source, consider how best to open a dialogue with the tax authorities about the quality of the data (bearing in mind the confidentiality of the data) and to monitor measurement error in the data, possibly through ad hoc studies, and ensure that the level of publication is appropriate for the quality of the statistics.
3. As part of the evaluation of VAT data, consider how defunct units will be identified and how businesses whose VAT data is simply late will be identified and the implications of any error in this identification for the estimates of movement in the PVI.
4. Initiate discussions within Statistics Sweden of the value of using a quarterly frame for the PVI, given the improvement in accuracy for short term movements this will bring.
5. Consider what the editing approach should be for the VAT data, as part of the new editing strategy.
6. In developing a regular analysis of revisions, use a graphical time series presentation, to make it easier for users to understand the trends, use a longer time series where possible, and develop analysis of the reasons for revisions.

Other areas for consideration

1. Continue to develop thinking about measures relevant to the accuracy of movement in the series, and the identification of turning points.
2. Continue with the work to evaluate the impact of the cut off of the smallest businesses.

Figure 1. PVI Ratings, Round 11

	Average Score Previous Round	Average Score Current Round	C1. Available Expertise	C2. Compliance with standards & best practices	C3. Knowledge of requirements, achievements, and improvement needs	C4. Plans for improvement activities	C5. Results of improvement activities and findings from other evaluations	C6. Communication with users and data suppliers	Importance to Overall accuracy (single sources of uncertainty)
Sub and sub-subcomponents of Accuracy									
Overall accuracy	35	40	○	●	●	○	●	●	
Sources of uncertainty:	34	40							
-Sampling	32	32	○	●	●	○	●	●	M
-Frame coverage	28	25	●	●	●	●	●	●	M
-Measurement	38	52	●	○	○	○	●	●	M
-Non-response	42	42	●	○	○	●	●	●	L
-Data processing	43	45	○	●	●	○	○	●	L
-Model assumptions	32	47	○	○	●	○	●	●	H
Preliminary statistics compared with final statistics	33	40	○	●	●	○	●	●	

Ratings					Importance to Overall accuracy and weights			
●	●	○	●	○	Not applicable (N/A)	Low (L)	Medium (M)	High (H)
Weak	Fair	Good	Very good	Excellent	0	1	2	3
1,2	3,4	5,6	7,8	9,10	Improvements		Deteriorations	

4.3. Quarterly Emissions Accounts (QEA)

Context and Discussion

Environmental Accounts aim to summarize and describe the impact on the environment from economic activities and development. The Quarterly Emissions Accounts is a subcomponent of Annual Emissions Accounts.

Users of air emissions statistics include Eurostat, policy analysts, researchers, and the Environment Protection Authority (EPA) who produce the greenhouse gas inventory. Politicians and the media are particularly interested in the quarterly statistics for their timeliness. A meeting is held with the EPA after every quarterly release, to exchange knowledge and discuss quality concerns. These discussions often lead to quality improvements, for example, the need for a residential adjustment was identified through user consultation, and this became a priority during the past year.

Model parameters and input data sources vary by industry and fuel type. Work to improve the estimates by refining model parameters and ensuring timely reception of input data is continuous. However, the work is prioritized based on the availability of new information to refine model parameters, which is not necessarily well aligned with the most significant contribution to overall estimates. This is noted not so much as a criticism, more as an acknowledgement of the challenges of working in this domain.

The production team have demonstrated understanding of the science of air emissions both in the quality report and in the ASPIRE interview. They are professional, curious, and strive for continuous improvement. They have built-in mechanisms to assure quality of their production work, for example internal documentation is kept up to date and the team rotate through production duties on a regular basis. The statistics are produced according to all available standards, e.g. Eurostat manual on air emissions accounts, and UNFCCC standard for emissions inventories. The production team was commissioned by Eurostat to write the latest edition of the manual for the Annual Air Emission Accounts which is regarded as the standard to follow even for the quarterly accounts (in the absence of a quarterly manual).

The production team members have a very good awareness of the general quality of the estimates. Since the previous ASPIRE review they have created a table in the quality report showing the share of total emissions contributed by each industry, and the proportion of input data for each industry that comes from survey data. This is informative, however the reader would need to be familiar with the quality of the input data sources to know the implications. For example, in the mining industry, 99% of the data comes from surveys, which sounds reassuring, however we do not know the quality of that data without consulting other sources. We also see in the table that mining accounts for only 2% of greenhouse gas emissions, so no matter the quality of the source, the overall impact is modest. On the other hand, agriculture, forestry and fisheries account for 18% of total greenhouse gas emissions (the second highest contributing industry), and only 1% of the data comes from surveys. Here the reader definitely needs more information to know if the estimates are fit for purpose. The production team are encouraged to make this table even more informative by ordering industries by their relative contribution to total emissions, and adding what they know about the quality of the input data, based on key sources and known strengths and weaknesses of those data sources. We also talked about having the production team pass along their level of confidence in the emission models used by EPA for each industry in the form of some sort of scale or rating. A simple star rating (1 to 5 stars) such as is used for restaurants would be immediately recognizable and interpretable.

The quality report included a list of revisions, including explanations of the source and reason for each revision and its effect on estimates. This is excellent, for transparency and to have a historical record. However, the impact over time and any comparisons through time or between levels is very difficult to see without a visual representation. We discussed plotting a time series of revisions at aggregate level from first publication, to final, for each estimate and doing analysis to identify trends, turning points, seasonality, etc.

While it is clear that the production team have an inherent feel for the quality of their statistics, they do not currently have methods for estimating either the accuracy or the uncertainty of the statistics. Granted this is not straightforward, given that the statistics are all model estimates. Estimating uncertainty would typically be in the skillset of a statistical analyst or methodologist, and methodology resources are very limited in this program.

The ASPIRE reviewers found reference papers from Denmark (Fauser et al, 2011) and the Netherlands (Ramírez et al, 2008) describing the application of a Monte Carlo simulation to estimate uncertainty in their greenhouse gas emissions inventories. We discussed the possibility of doing a pilot study using Statistics Sweden's air emissions data. This would be an opportunity for advancement on several fronts. Most practically, this method would result in estimates of uncertainty for the published air emissions statistics, which would be of value to all data users. The inclusion of methods in the manual on producing air emissions statistics that has been provided to Eurostat would mean that others producing these statistics would also be able to include estimates of uncertainty. These would be clear and highly visible examples of Statistics Sweden demonstrating statistical leadership. Also, this would serve as a learning opportunity for methodologists at Statistics Sweden, who could then potentially find other applications for the method and thereby improve knowledge of quality of other products.

Key recommendations for the coming two years

1. Improve the usefulness of Appendix 1: Table of uncertainty by making the following modifications:

- a) Order industries from greatest to least contribution to total emissions
- b) Include text descriptions and hyperlinks to other reference material (where appropriate) describing what is known about the quality of data sources, whether they are survey data or non-survey data; include quantitative measures where they are available (this would only need to be done for key data sources for each industry and relate to important quality issues for these sources.)
- c) Add some sort of rating, for example stars (1 star is low, 5 stars is high) indicating the confidence that the production team have on the emission models used in each industry, based on the assessed reliability of the emissions models used.

2. Perform 'revisions analysis', which would include plotting final and revised estimates through time. Analyze the size and direction of revisions; look for trends, turning points, seasonality, correlation, etc. Compare what you see in the graphs to what has actually happened in industry; does it make sense? Share and discuss your analysis with data users.

3. Initiate a research project for methodologists to explore the possibility of using a Monte Carlo simulation to estimate uncertainty of model estimates. They will require time to do this work, possibly a short contract with a university student to do computer programming work, and they will need data to work with, possibly quarterly, annual or even inventory data. We would encourage the Statistics Sweden methodologists to reach out to their counterparts at other statistical agencies who have published similar work (Netherlands, Denmark, Austria). They should also be encouraged to consult, collaborate and share findings with their colleagues at Statistics Sweden, who might find other applications for this method.

Other areas for consideration

1. While we fully acknowledge and applaud the production team members for their knowledge and expertise in air emissions statistics, we see a gap in knowledge of statistical methods to estimate accuracy and uncertainty. It's a classic "they don't know what they don't know" scenario, which we have seen often in the form of under-involvement of methodologists in statistical programs. A methodologist with more involvement in analytical discussions and strategic decisions could bring a different perspective and skill set which would complement those of the subject matter experts.

Figure 2. QEA Ratings, Round 11

	Average Score Previous Round	Average Score Current Round	C1. Available Expertise	C2. Compliance with standards & best practices	C3. Knowledge of requirements, achievements, and improvement needs	C4. Plans for improvement activities	C5. Results of improvement activities and findings from other evaluations	C6. Communication with users and data suppliers	Importance to Overall accuracy (single sources of uncertainty)
Sub and sub-subcomponents of Accuracy									
Overall accuracy	42	47	○	○	○	▲	○	▲	
Sources of uncertainty:	42	47							
-Sampling	40	40	○	■	▲	●	●	○	L
-Frame coverage	42	53	○	○	○	○	○	○	M
-Measurement	42	42	■	■	▲	●	●	○	L
-Non-response	42	42	■	■	▲	●	●	○	L
-Data processing	48	50	■	■	○	○	▲	●	L
-Model assumptions	42	48	○	■	▲	▲	▲	○	H
Preliminary statistics compared with final statistics	32	33	▲	▲	▲	▲	▲	▲	

Ratings					Importance to Overall accuracy and weights			
●	▲	○	■	●	Not applicable (N/A)	Low (L)	Medium (M)	High (H)
Weak	Fair	Good	Very good	Excellent	0	1	2	3
1,2	3,4	5,6	7,8	9,10	Improvements		Deteriorations	

4.4. Building Statistics

Context and Discussion

The Building Statistics series evaluated in ASPIRE Round 11 includes the three collections: Permits, residential and non-residential, and Commencements and Completions, residential only.

The Swedish National Board of Housing is a key user of the statistics and is also expressing interest in new statistics on Commencement and Completion data for non-residential building. The data, particularly the Commencements and Completions data, is used by the National Board as input to their forecasting models. The monthly Permits data is also used by the National Board and the Riksbank as an economic indicator in their forecasting. While Statistics Sweden only publishes the quarterly permits data, and in a non-seasonally adjusted form, for the general public to use, it provides the monthly permits data to the National Board of Housing which also provides it to the Riksbank. Statistics Sweden also provides seasonally adjusted monthly permit data to Eurostat where it is published as an economic indicator series.

While Statistics Sweden has a good relationship with the National Board of Housing and with Eurostat, there continues to be less communication with other users. The potential use of monthly permits, preferably seasonally adjusted and smoothed, has not been discussed with other analysts or users in Sweden, who might value such series as an early indicator of building activity, and economic activity more generally.

There appears to be an opportunity here for Statistics Sweden to play a statistical leadership role and build their own understanding of the value of the monthly seasonally adjusted permits data as a leading indicator of economic activity, and if it proves valuable, to publish it for general use to inform discussion about current economic conditions. This is done in a number of countries and leads to the first recommendation below.

A significant quality issue for the preliminary building statistics is the lag in reporting by municipalities which can be of the order of 30 to 40% for the quarterly number of building permits and is still around 10% after 4 months. The lag in permits affects the commencements and completions data as well since the permits data provides the frame for these collections. An adjustment is made for the lag in reported statistics, with the adjustment being based on an average number of missing permits for the same period over the previous three years. New work has been done to evaluate the adjusted preliminary data compared to the final data, indicating the adjustment performs quite well over recent years. This work is commended. It is intended to repeat the study already undertaken on approvals data, on the commencements and completions data to identify the impact of revisions on these series.

This further work is encouraged, and in addition it is recommended that a further, simple analysis be undertaken that plots the final data against the adjusted preliminary data as first published, over time to highlight the performance of the adjustment during periods of upturn and downturn in economic activity. It is also recommended that a presentation such as this plot be added to the Quality Report to inform users of the accuracy associated with the adjustments. As more information on quality of the building statistics is gained, it is important for the Quality Report to be kept up to date.

The high level of lag in the collected data is largely the result of complex and costly collection mechanisms involving significant manual intervention and querying between Statistics Sweden and the municipalities. While there is every intention to move to a more streamlined collection and editing approach, in line with the Statistics Sweden's drive in this direction, a timeline for such a transition is not yet in place. Mapping of processes has occurred leading to improved understanding of what is needed and identifying significant inefficiencies with the current approach. While some automation of edits, imputations and derivations has occurred, scope for substantial improvement has been identified. The ability to move forward on this is limited by the fragility of the current system. In the meantime, the resources of the project team are being used to supplement the data collection effort in order to simply produce required statistics on time. This is inefficient and frustrating for all involved.

The solution to the systems issue seems to involve three factors. Firstly, work must be done with the municipalities to get engagement for better processes with current systems in the short term, and to build their buy-in for machine to machine (M2M) data transfer in the longer term. A second body of work relates

to identifying and implementing the automation that is possible within Statistics Sweden, with the newly implemented files-based approach to data collection. Finally, a time frame for the proposed system redevelopment based on M2M and the new editing approach needs to be identified to allow planning of the work program to be started so that the first two streams can proceed in a way that works, effectively towards the longer-term solution.

If the data collection system is improved, the response patterns of municipalities will change, impacting the lag and hence lag adjustment. How this transition period will be managed in terms of modelling the adjustment needs to be considered. A past recommendation from ASPIRE was to understand response histories of the municipalities. This would allow for targeted follow up in data capture and provide a basis for problem resolution and engaging with municipalities regarding introducing M2M. This recommendation is repeated.

As well, a top-down building of relationships with each municipality, recognising them both as important users as well as providers of information, might facilitate productive cooperation. A client manager in Statistics Sweden, working across the municipalities and arranging user contact meetings on different topics, as well as discussions on improving processes for data provision, might help build these relationships.

Key recommendations for the coming two years

1. Statistics Sweden set up a joint project involving analysts, the Building statistics team and methodologists to develop an understanding of the value of building approvals, commencements, and completions as economic indicator series, for general use to inform discussion on current economic conditions in Sweden. If a decision is made to progress publication of the data, appropriate presentation should be considered to optimise the communication of the information within the series (e.g. graphs of seasonally adjusted, and smoothed seasonally adjusted series).
2. Extend and communicate to users, the evaluation of the lag adjustment.
3. The Building team continue improving data capture and editing as far as is possible within the current system, while reaching corporate agreement for a longer-term redevelopment of the front end of their system.
4. The Building team work with other corporate areas to build a strategic relationship with municipalities with regard to both the use and collection of statistics, with a view to establishing effective M2M interfaces for data collection.

Figure 3. Building Statistics Ratings, Round 11

	Average Score Previous Round	Average Score Current Round	C1. Available Expertise	C2. Compliance with standards & best practices	C3. Knowledge of requirements, achievements, and improvement needs	C4. Plans for improvement activities	C5. Results of improvement activities and findings from other evaluations	C6. Communication with users and data suppliers	Importance to Overall accuracy (single sources of uncertainty)
Sub and sub-subcomponents of Accuracy									
Overall accuracy	28	35	●	●	●	●	●	●	
Sources of uncertainty:	34	37							
-Sampling	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
-Frame coverage	27	35	●	○	●	●	●	●	H
-Measurement	40	38	●	○	○	●	●	●	H
-Non-response	35	35	○	●	●	●	●	●	L
-Data processing	40	40	○	○	●	●	●	●	M
-Model assumptions	20	38	○	○	○	●	●	●	L
Preliminary statistics compared with final statistics	27	35	○	●	●	●	●	○	

Ratings					Importance to Overall accuracy and weights			
●	●	○	●	○	Not applicable (N/A)	Low (L)	Medium (M)	High (H)
Weak	Fair	Good	Very good	Excellent	0	1	2	3
1,2	3,4	5,6	7,8	9,10	Improvements		Deteriorations	

4.5. Population statistics

Context and Discussion

The Population Statistics describe the size and composition of the Swedish population, and the component series of births, deaths and marriages, as well as statistics on households and on asylum seekers. Both national and regional population numbers are provided as well as statistics cross classified by core demographic variables.

The Population Statistics area maintains an ongoing relationship with users through the User Board which includes demographic analysts, modellers looking to forecast for planning purposes at municipality and county level, and users with an interest in migrant settlement and integration. The User Board appears generally content with the quality of the statistics and discussions are largely driven by Statistics Sweden, although the academic analysts will also raise issues in demographic trends appearing in the data for discussion. There appears to be little discussion about, or concern expressed in relation to quality issues, other than more recently the lags in flows of migrant data from Tax through to the TPR.

The population statistics at the total population level, appear well developed, and well documented, with a large amount of useful quality measurement reported. Over-coverage is modelled and updated regularly. Under-coverage due to lags in the registrations flowing from Tax Agency, is well understood, and recent increases have led to planned improvements to the Tax registration process. Under-coverage due to failure to register or seek asylum (undocumented migrants) is less easy to measure but analysis of potential sources has indicated the number involved is small. Continued investigation of alternative sources to measure this aspect of under-coverage is encouraged.

In addition to population totals for all of Sweden, the Total Population Register is a very rich and timely source of information on population characteristics for lower levels of geography and for various socio-demographic subgroups. Population estimates are provided at the municipality and county level and for subgroups based on age, sex, and country of birth and citizenship. The TPR is also linkable to other registries for more detailed analytic work.

It is of concern to the ASPIRE team that, while the richness of the TPR lies in its ability to provide population data at the regional level and for various subgroups, there is very little evaluation of population statistics below the total population level. For example, while over-coverage studies provide regularly updated quality measures at the total population level, no such measures are analysed or available for regions or sub populations, even though that would be relatively easy to do, and even though it is recognised that some groups will be significantly more affected than others.

Similarly, the quality report refers to the possibility of cross-checking aspects of the TPR quality using comparison with other registers, however such studies are not reported although they too could provide analysis at the regional and subgroup level where particular problems might be clustered.

Population estimates for subgroups of the national total will also be of interest to some users as the denominators, when rates of certain characteristics are compared between different groups for example in crime and justice statistics, or mortality rates. The quality of the denominators, for example in terms of over or under-coverage, may impact the validity of comparisons made. This is another reason to provide quality information for key regional and subpopulation level estimates.

The TPR also provides estimates of households and type of household. These are based on individuals with shared addresses and linkage with the Dwelling Register. Imputation for addresses not found in the Dwelling Register has improved household counts, so that it is expected that counts of household and type of household are more accurate and will conform more closely with the current census than they did with 2011 data, where significant quality issues were found. The current census provides an opportunity to evaluate the improvement made. However, prioritisation of resources means this evaluation will not be performed.

While there is close attention paid to population movements into and out of the country, and the accuracy of these movements, the impact of unregistered moves within Sweden on county and municipality estimates is not evaluated. No measures are provided in the Quality Report of the occurrence of incorrect address on the register, although it is known that some groups, such as grown children leaving home, have higher

levels of incorrect address. These errors will not only affect population estimates, but also household estimates, and type of household estimates.

An independent source of quality measurement for households and for addresses of individuals within Sweden, may be the household survey program, with selections based on the TPR. Feedback from such surveys could provide an indication of the quality of the register in terms of core variables, including occupants of a given address, and household information.

While such quality studies require resources, the running of a population census in Sweden provides an opportunity to design a joint evaluation program for both the census and the population estimates as subnational level drawn from the TPR. The census will need to evaluate quality in order to describe the results. The evaluation program could look at various options for understanding quality including confrontation with alternative register sources, and also estimates based on field findings from household surveys, as feedback to both the census and the TPR.

While Swedish users do not appear to have a concern about the current conceptual basis for measuring permanent population in Sweden, the EU is proposing a new definition involving the concept of 'intention to stay 12 months'. This will raise some measurement issues for particular population groups, for example those Ukrainian migrants on temporary migrant visas limited to 11 months and 30 days, who may well intend to reapply. The User Board appears an appropriate mechanism for working through some of the user implications for such changes. The Population Statistics area is developing a proposal to support such a change in definition through the development of the Individual Base Register (IBR) which would include the additional population present in Sweden but not on the Swedish Population Register. If this approach is taken this will be a major new development for the area.

Key recommendations for the coming two years

1. We commend the work that is done on quality at the total population level but given the importance of population estimates at the regional level and for particular population groups, quality analysis also needs to be undertaken at this level and the results analysed, documented and discussed with users of this data. In particular, given that a census is currently being undertaken, we encourage a joint evaluation program to provide information on both the census results and the quality of population estimates based on the core TPR variables such as age, sex, marital status, country of origin as well as household and household type information.
2. To better understand the impact of incorrect addresses (e.g. as a result of failure to register change of address) on regional population estimates, strategies to estimate and analyse the incidence of incorrect addresses within Sweden, on the TPR should be explored. Possible sources of analysis include data confrontation between TPR and additional registers. Another approach is to reconcile data on the TPR with information from contacts with households as part of the household survey program. The household survey program might also provide some information on the quality of household data within the TPR.
3. The user consultation at Board level should be supplemented with feedback from those undertaking analysis of core data in the TPR. In particular, feedback loops should be strengthened, on quality issues arising from other areas of Statistics Sweden involved in providing TPR based outputs to users, for example on commission.
4. The Population Statistics team intends to build in house training, to help minimise the impact of staff turnover. This is strongly encouraged, and we would also encourage the team to build their analytic skills in terms of understanding and being able to effectively analyse and communicate on quality aspects of the register and how these impact population estimates at the sub population group and regional level.
5. As part of building analytic capability, the level of methodological resource applied to the program of statistics should also be reviewed as it is currently set at a very low level given the significance of the TPR to understanding Swedish population data.

Figure 4. Population Statistics Ratings, Round 11

	Average Score Current Round	C1. Available Expertise	C2. Compliance with standards & best practices	C3. Knowledge of requirements, achievements, and improvement needs	C4. Plans for improvement activities	C5. Results of improvement activities and findings from other evaluations	C6. Communication with users and data suppliers	Importance to Overall accuracy (single sources of uncertainty)
Sub and sub-subcomponents of Accuracy								
Overall accuracy	43	○	○	▲	▲	▲	▲	
Sources of uncertainty:	43							
-Sampling	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
-Frame coverage	42	○	○	○	▲	●	○	H
-Measurement	40	▲	▲	▲	▲	▲	○	M
-Non-response	47	○	○	○	▲	▲	○	L
-Data processing	48	○	○	▲	○	▲	○	L
-Model assumptions	43	○	○	▲	▲	▲	▲	H
Preliminary statistics compared with final statistics	40	○	▲	▲	●	▲	○	

Ratings					Importance to Overall accuracy			
●	▲	○	▼	○	Not applicable (N/A)	Low (L)	Medium (M)	High (H)
Weak	Fair	Good	Very good	Excellent	Weights			
1, 2	3, 4	5, 6	7, 8	9, 10	0	1	2	3

5. References

- Fausser, P., P. Sørensen, M. Nielsen, M. Winther, Marlene S. Plejdrup, L. Hoffmann, S. Gyldenkærne, M. Hjorth Mikkelsen, R. Albrektsen, E. Lyck, M. Thomsen, K. Hjelgaard and O. Kenneth Nielsen. 2011. [“Monte Carlo \(Tier 2\) uncertainty analysis of Danish Greenhouse gas emission inventory”](#). *Greenhouse Gas Measurement and Management*, Volume 1, Issue 3-4: 145-160.
- Linacre, Susan, Stephen Penneck and Laurie Reedman. 2021. [A Tenth Application of the ASPIRE Quality Evaluation System for Statistics Sweden](#). Accessed September 14, 2022.
- Statistics Sweden. 2021. A System for Product Improvement, Review and Evaluation, ASPIRE – a manual. Accessed September 14, 2022. [A System for Product Improvement, Review and Evaluation, ASPIRE - a manual \(scb.se\)](#)
- Ramírez, Andrea, Corryde Keizer, Jeroen P. Van der Sluijs, Jos Olivier and Laurens Brandes. 2008. [“Monte Carlo analysis of uncertainties in the Netherlands greenhouse gas emission inventory for 1990–2004”](#) *Atmospheric Environment*. Volume 42, Issue 35: 8263-8272.