A Tenth Application of ASPIRE for Statistics Sweden

Susan Linacre, Stephen Penneck, Laurie Reedman

June 16, 2021

Contents

1. Executive Summary

Four previous products were reviewed in 2021: Consumer Price Index (CPI), Producer and Import Price Index (PPI), Statistical Business Register (SBR) and Quarterly Gross Domestic Product (GDP-Q). These products were last reviewed in 2019. Section 2 gives the background to ASPIRE, and sets out the changes in this ASPIRE round from previous rounds. The key findings are in section 3, and are summarised below.

All product areas showed a high level of professionalism in relation to their products and were constructive and supportive despite the difficulties imposed by the remote working arrangements.

The new organisational structure is an important driver for change, and, supported by a deeper cultural change programme, will help build cross-cutting working, broader team working and collective ownership of strategic issues. The role of the new Methodology and Architecture Management Unit, reporting directly to the Director General, will help continue to build methodological capacity while supporting a more corporate approach to product development. The methodological capability of a national statistical institute underpins the rigour and hence reputation of its statistical products. It also supports the rationalisation and integration of statistical processes as needed to fit a more corporate approach to product development and maintenance. Strong statistical institutes typically have a strong methodological capability, with senior representation in corporate decision-making fora.

Developing statistical leadership throughout the organisation will be key to success, and we hope our setting out more clearly in section 3.1 our understanding of what this means for statisticians will be helpful.

Our review took place over a year since the beginning of the COVID pandemic, which has shown the vital importance of relevant and timely statistics for government. Statistics Sweden has developed new products at pace, and will want to reflect on how this experience can help them move to the more responsive agency they seek.

We have been impressed by the response of the agency to the challenges to statistical quality that the pandemic has brought. The pandemic has changed consumer and business behaviour, and Statistics Sweden will need to be mindful of how this is reflected in their statistics as the economy continues to recover, providing clear explanation of the quality implications for users.

As with many statistical offices the demand for IT services outstrips supply. Moving further towards the use of common tools will help here, with some compromises needed in product areas to minimise the need for adaptations.

The Consumer Price Index has a strong development programme, and is also fully challenged by the changes in consumer behaviour resulting from the pandemic. This will all add to the risk for the product over the coming period.

The focus of the Producer and Import Price Index has been on improving the coverage for services. An attempt to develop implicit quality indices, which provide users with a view of the impact of quality adjustments, was abandoned. We have suggested that the feasibility of progressing this work should be studied.

The Statistical Business Register is a significant part of the statistical infrastructure with the ability to provide coherence to economic statistics. An up-to-date register adds to timeliness, and it would be worthwhile to review the relevant processes to ensure this benefit is fully realised.

Quarterly GDP has also seen significant development, with the publication of flash estimates. We are encouraging the completion of work to incorporate the results of the quarterly Structural Business Survey, at the earliest opportunity.

Between the previous review of these products in 2019 and the current review, the ASPIRE checklists have been modified (see description of changes in Section 2). These changes make comparisons between 2019 and 2021 less meaningful. However it is worth noting that the overall accuracy score for the SBR has

increased, reflecting the high level of corporate attention given to the SBR over the past 2 years. Good progress has been made towards many of the recommendations from 2019, as described in section 3.

Some of the issues summarised above, together with others, are set out section 4, which covers cross-cutting issues and recommendations. In summary these are:

1. Use the opportunity of the new organisational structure to encourage statistical leadership at all levels of the organisation.

2. Use the experience of the COVID-19 pandemic as a catalyst to becoming more responsive to emerging issues and to further improving communication with data users.

3. Examine the impediments to updating frames for sub-annual surveys quarterly, and consider the benefits of moving towards common frames for sub-annual surveys rather than the current practice of survey areas validating and maintaining their own frames.

4. To alleviate the risk posed by a shortage of IT resources, continue to move towards providing corporate functionality while ensuring each product area is able to work effectively with the corporate tools.

5. To stay abreast of the data revolution, further develop expertise in data science and machine learning techniques.

6. Consider what options there are for collecting household spending data that could reach an acceptable level of quality.

2. Introduction and changes to ASPIRE in Round 10

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.

Following a review of the process in 2018, some changes were implemented in 2019, and these were further refined for 2020. Five of the ten previous products were rotated out in order to provide space for new products which were 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.

While the general ASPIRE process has remained similar through the years, the 2020 and 2021 review rounds differed with more specific checklists used in scoring and the remote assessment due to the COVID-19 pandemic.

Changes to the checklists
The new external review team engaged in 2018 requested that the rating process be made more comprehensible, transparent and reliable both for the experts themselves and for the product teams. In collaboration between Statistics Sweden and the review team, clarifications were therefore made in the checklists which support the rating process. Briefly, this involved:

- renaming three of the six criteria to better reflect the cyclical process of improving quality in relation to quality requirements such that the statistics be fit for purpose,
- breaking down the checklist requirements into the same number of levels as there are ratings i.e. 10 levels,
- specifying in more detail what is required for each of the levels.

The above changes seemed effective in improving the rating process as needed.

A detailed description of the changes and present checklists can be found in the ASPIRE-manual (Statistics Sweden 2021).

Remote assessment
Due to the prevailing circumstances with the COVID-19 pandemic a remote assessment was performed between the review team and the product teams, similar to the 2020 review.

The product teams did their necessary preparatory work with the checklists and quality documentation which was sent to the expert team as usual. Thereafter, the review team had a written exchange with follow-up questions and answers with each of the product teams. An online meeting was held with each product team and the review team to resolve any outstanding questions and share the preliminary results.

Given the circumstances, the remote set-up worked somewhat better than in 2020. In 2021 there was only one round with a written exchange instead of two rounds in 2020. One round proved sufficient.

Changes in the external review team
The external review team was made up of three experts – Susan Linacre and Stephen Penneck who joined the team in 2018 and Laurie Reedman who joined in 2021. The team was thereby increased by one expert compared to round 9 in 2020.

The review team would like to express their gratitude to staff in Statistics Sweden who have participated in ASPIRE round 10 at this difficult time. We were conscious of the additional challenges that the COVID-19 pandemic makes on their time, and also the different operating circumstances. We are also grateful for the quick responses we got to our questions and for the open way that staff responded.

Nonetheless we recognise that conducting such a review at distance is far from ideal. We hope we have understood the quality issues facing these product areas, but it must be recognised that this review has had more limitations than is usual.

3. Product Reviews

3.1 General Observations

We were briefed on the new organisational structure, which is an important driver for change in Statistics Sweden. We understand that reducing the number of departments and setting up a new data department will create a more dynamic organisation with the ability to adapt and remain relevant in the face of the data revolution that is affecting government and society.

We can see that these changes, supported by a deeper cultural change programme, will help build cross-cutting working across Statistics Sweden, supporting broader team working and collective ownership of strategic problems.

The role of the Methodology and Architecture Management Unit reporting directly to the Director General will be particularly important in building a changed culture and capability in the organisation if the approach taken is to build agility through effective infrastructure with shared tools to perform common business processes. The central Methodology unit will have an important role in continuing to build the methodological capability in the organisation both centrally, in the two statistics departments and in the data department. The methodological capability of a national statistical institute underpins the rigour and hence reputation of its statistical products. It also supports the rationalisation and integration of statistical processes as needed to fit a more corporate approach to product development and maintenance. Strong statistical institutes typically have a strong methodological capability, with senior representation in corporate decision making fora.

While the statistical leadership of methodologists will be important, developing statistical leadership throughout the organisation will be key to the success of the needed culture change across the organisation. Statisticians in a National Statistical Office are professionals, with expertise in data, its collection, and analysis, and also in understanding its quality attributes and the implications of these for the key uses of the data. They have a professional set of values associated with ensuring the integrity of statistics and have a duty of care to their users. To be effective they must show statistical leadership in their areas of expertise. This leadership extends outside the organisation to the users, with whom there must be strong engagement, and also inside the organisation to areas of specialist expertise who must be partners in assuring that best practice approaches are being used across the agency, and to areas who produce any source data that they use.

By statistical leadership we mean statisticians:

- using their understanding of their statistical product and its quality attributes to engage strongly with users on the best possible uses;
- developing with colleagues clear goals for the part of the organisation they manage consistent with their role, and the office's goals and values;
- developing and delivering a clear direction for their product to meet perceived needs;
- helping colleagues to deliver this vision, by empowering them, helping them overcome obstacles and take managed risks;
- collaborating with colleagues to ensure the product remains relevant in a changing world;
- and contributing to corporate initiatives and infrastructure and helping the development of best practice approaches.

Our review took place over a year after the beginning of the COVID-19 pandemic, which has been an acute test of Statistics Sweden's ability to react quickly to a rapidly changing environment. Statistics have been critical to governments' approach to this challenge, and this has been a test for the statistical leadership of all statistical offices. At a time when working conditions have been disrupted, offices have had to develop

new data collections and analyses at pace, and change collection and processing approaches to meet changing circumstances. The crisis has shown what statistics offices can do under duress to meet vital user needs, and Statistics Sweden will want to reflect on how this experience can help them adapt to the vision of a more responsive agency they are now pursuing.

We were interested in the impact of COVID-19 on the quality of the statistics we were reviewing. Clearly there was an impact on survey response rates, though we were impressed with the extent to which rates for business surveys had held up. In addition, consumer and business behaviour changed, and the assumptions in the compilation of key statistics such as GDP and the CPI needed to be revised (with very little information). Behaviour will change again as the Swedish economy continues to recover. These changes will add to the uncertainty of the statistics. Coming at such a key time it becomes even more important that there is very good communication with users on these issues, so that, in the event of future revisions, users are aware of some of the difficulties in making good estimates at this time, and feel they had an opportunity to contribute to solutions where appropriate.

As Statistics Sweden seeks to gain agility as well as consistency and quality, it will rely on cross organisational infrastructure, for example such as we understand will be developed in the data department. It is important that the benefits of this common infrastructure are reaped through consistent use across the organisation.

The Statistical Business Register is a significant piece of organisational infrastructure that provides a very current view of the business population and its key attributes. However survey areas, individually, decide whether to use the quarterly snap shots of the SBR to update business survey frames each quarter. Currently sub-annual surveys choose not to update their frames each quarter because of the tailored work involved in preparing the frame and selections for their particular view. To reap the benefits of having such an up-to-date register, it would be worthwhile to review the processes associated with a survey area using an SBR snap shot, and refreshing the sample on a quarterly basis. What processes could be rationalised and done once across all survey areas if centralised? What information is being held in separate business areas about population units that could be shared in a common frame? The costs and benefits of the current survey by survey approach to population and sample management should be considered against the option of a more centralised approach.

More generally, the steady demand for "traditional" IT resources, that is, to build, test, support and maintain production systems, is outpacing the supply. Moving from each program having its own set of production systems to shared use of common tools has already begun and should continue. This will pay off in reduced maintenance costs and an increased agility in taking up new technical opportunities as they arise. This will require some compromises, with product areas changing some of their methods to adapt to those supported by common tools. The benefit of each case will need to be assessed. This will provide functionality to ensure each product area is able to work effectively with the corporate tools. Common tools and systems will alleviate the current problem of competing demands and reduce the risk of quality posed by a shortage of IT resources.

The data revolution has sparked a growing need for data science skills, in particular machine learning techniques to do exploratory work, and data analysis on large amounts of data, for example web scraping, analysing scanner data, interpreting satellite images of crop land, coding and classification. We encourage Statistics Sweden to develop further expertise in data science and machine learning.

Declining household survey response rates is a worldwide problem. In Sweden this resulted in the Household Budget Survey being discontinued ten years ago. As well as being a key source of data for consumers' expenditure (a key component of GDP), the survey also provided the high level weights for the CPI: thus contributing to two key economic indicators. Alternative sources have been found for the national accounts – essentially using turnover data. Work is also progressing to find sources for the CPI – namely

scanner data and business survey data. However none of these sources directly measure what consumers spend, just what business sell; and the gap between them is difficult to measure – especially when more consumer spending takes place over the internet, with an increasing proportion coming from overseas.

3.2 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 10, this would involve comparing scores with those for round 8 when they were reviewed last. However, such comparisons should be made considering the clarifications made in the checklists prior to round 9 in 2020 and described in section 2 – clarifications which had the effect of tightening up requirements to achieve a higher score. This effect could be noted for some products particularly regarding Criteria 1 *Available expertise* and C2 *Compliance to standards and best practices*, where a lower score was assigned without observing a real deterioration in the quality of the product, compared to round 8. Nevertheless, the break in the time series is assessed to be within the margin of error spoken of above and comparisons over time within a product on an overall level should still indicate if the product is improving the accuracy in statistics or not.

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)
CPI	58	64
PPI	55	56
SBR	63	61
GDPQ	58	57

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 for statistics/registers	CPI	PPI	SBR	GDP-Q	Average score	Weighted average score
Overall Accuracy	58	55	63	58	59	N/A
Sources of uncertainty:	64	56	61	57	60	N/A
-Sampling	70	55	N/A	47	43	42
-Frame coverage /Coverage	63	60	62	52	59	54
-Measurement	73	58	62	65	65	52
-Non-response	55	52	40	47	49	28
-Data processing/Processing	57	55	65	55	58	49
-Model assumptions	62	53	63	62	60	51
Preliminary statistics/register compared to final statistics/register	N/A	N/A	N/A	63	N/A	N/A

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

3.3 Product Ratings and Recommendations

3.3.1. Consumer Price Index, CPI

Context and discussion

The CPI is the key measure of inflation, used by government and more widely by businesses and citizens. It is used for economic policy, as a deflator in the national accounts, and for indexation purposes.

The Swedish CPI continues to be of very high standard, especially when compared with other countries. 2020 was dominated by responding to the COVID pandemic, which impacted on consumer and business behaviour and had major implications for the compilation of the CPI. The Prices team engaged internationally on this issue and responded well to the challenge. The issue, and the diversion of the team from other quality improvement projects, will have added to the uncertainty around the CPI at this time.

There is a strong development programme, taking in new data sources and innovative methods. However, the scale of these strategically important projects, including the introduction of replacement weights for the HBS, the replacement of survey data by scanner and web scraped data, and the introduction of a new data processing system come on top of all the regular changes that need to be made to maintain the CPI as fit for purpose, and the need to ensure the index fully reflects the short term changes to the Swedish economy caused by the COVID pandemic and will add to the risk for this product over the coming period.

Progress towards prior recommendations

1. *CPI weights:* In 2019 we recommended that work on alternative sources to the Household Budget Survey should continue and be given more emphasis, and the sensitivity of the CPI to changes in the weights should be assessed.

   We were pleased to hear that a highly prioritised (though small scale) project to replace the HBS, including using scanner data and data from the Structural Business Survey was underway, although resource limitations had prevented an assessment of the sensitivity of the CPI to weights changes.

2. *Data processing of scanner data:* In our previous report we recommended that given the significant use of scanner data in the CPI operations, consideration should be given to developing an office-standard IT processing system to replace the development systems in place and reduce risk.

   We were concerned to hear that resource pressures had prevented this from being taken forward. However the new strategic approach (NDK) could provide a solution for this.

3. *Statistical estimation from large data sets:* We recommended that consideration should be given to how the large data sets now being collected can best be distilled into price indicators, taking into account the price behaviour of consumers.

   A Eurostat funded project, currently underway, will recommend an improved methodology which will reduce total survey error.

Other accomplishments

1. There has been good use of scanner, API and web scraped data, which now make up 45% of the CPI and is increasing. This has required improved data processing, and new checking and analysis procedures.

2. Changes to data sources, samples, imputation methods and model assumptions have been made to reflect the changing pattern of consumer behaviour during the pandemic.

3. Additional meetings have been held with users and an open approach has ensured good understanding by them of the impact of the pandemic.

4. There continues to be a research programme – e.g. on the impact of modern consumer behaviour – as well as a programme of evaluation of methods and implementation of new methods for products – e.g. the method for Monthly Chaining, overlap methods, and methods for pharmaceuticals, tenant-owned dwellings.

5. The quality of source data used for weights is now reviewed annually and used to prioritise developments.

6. Internal guidelines have now been developed to ensure quality is maintained when implementing new data or methods.

Key recommendations for the coming two years

1. *Replacement source for the Household Budget Survey.* Work has continued on using turnover data (such as scanner data) and the Structural Business Survey as alternative sources, on a small scale. Weights for some areas of the CPI (e.g. clothing and furnishing) remain out of date. Implementation of this work will not be until 2023, and this work should continue to be given priority. There will be difficulties in finding appropriate sources in all cases, such as sales from overseas, and problems in modelling household expenditure estimates from these sources. Resources will need to be deployed in these areas, and the results of the new approaches evaluated, including an analysis of the potential impact on the weights, to ensure the resulting revisions to the CPI weighting methodology are justified.

2. *Data processing of scanner data*. The current system is described as 'good enough' but it is not robust nor efficient. Given its evolution from development systems and the complexity of the various operations performed it should be considered an organisational risk. The NDK project should provide a strategic solution, but will need to be prioritised, resourced, timetabled and developed in close collaboration with the CPI team. The team should also investigate the use of machine learning techniques for investigative work with scanner data.

3. *Monitoring the impact of COVID.* COVID has led to changing patterns of consumers' expenditure and in changes to prices. It will be necessary to continue to monitor its effects, as the Swedish economy continues to recover and discuss them with users.

4. *Managing risk.* The introduction of replacement weights for the HBS, the replacement of survey data by scanner and web scraped data, and the introduction of a new data processing system will all come on top of all the regular changes that need to be made to maintain the CPI as fit for purpose. This will need careful risk management to ensure the impact on the CPI is known and well planned.

Other areas for consideration

1. Learning more from best practice outside the EU. Sweden is actively involved in EU Task Forces and Working Groups and in receipt of grant funding for projects. Given that all this is within the constraints of the Harmonised Index of Consumer Prices, the team could consider developing further the already good links they have outside Europe with the Ottawa City Group on Price Statistics to ensure they benefit from a wider range of peer discussion.

2. Researching the impact of growing sales from overseas. This is a particular weakness of the reliance on domestic sales weights for the CPI, given the growth in this area, and the potential for price development for overseas purchases to differ from domestic price development.

Figure 1. CPI Ratings, Round 10

Sub and sub-subcomponents of Accuracy	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)
Overall accuracy	58	58		○	○	○	○	○	
Sources of uncertainty:	63	64							
-Sampling	67	70				○	○	○	H
-Frame coverage	65	63				○	○	○	M
-Measurement	68	73							H
-Non-response	55	55		○	○				L
-Data processing	60	57	○	○	○	○	○		H
-Model assumptions	58	62		○	○		○	○	H
Preliminary statistics compared with final statistics	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	

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

3.3.2. Producer and Import Price Index, PPI

Context and discussion

The PPI measures the average price development for producer and import prices, in total and for different product groups. The PPI provides input to the National Accounts when calculating GDP in constant prices. PPI also makes an important contribution to evaluating current monetary policy.

The impact of COVID-19 on the quality of the PPI over 2020 was judged to be low, with non-response and the number of inactive units surveyed during 2020 being similar to 2019, with most change in non-active units occurring amongst those affected by pandemic restrictions, such as hotels, travel services. No changes were deemed necessary for the index in term of the methods used, compared with those used prior to the pandemic.

The focus of the statistical development for PPI since the last ASPIRE work has been on improving coverage and measurement in services.

A major development has been to move the index calculation method for services from geometric to arithmetic at the lowest level of aggregation, as of the first quarter of 2020. This was evaluated with back-data and shown to provide stable estimates. It brings the aggregation method for services into line with that for other industry sectors.

As of the publishing of the first quarter of 2021 product groups in service areas are now being published for different markets (domestic, exports and imports respectively), where possible, as well as the overall service producer price index (domestic + exports). The number of product groups published has increased from around 100 to 140 as of the latest publishing.

The wholesale and retail trade area is large in volume and in value added, but is a difficult industry for several reasons (including valuation basis; negative prices; weight sources and application; quality adjustment). The team has focussed initially on wholesale and retail trade for motor vehicles as a focal point to test theories and new methods that later have the potential to be scaled-up for divisions 46 and 47 – Wholesale and Retail trade, except of motor vehicles and motorcycles.

As part of a Eurostat funded Grants project, detailed research reports were written on G45 - Trade and repair of motor vehicles; H52.21, 52.22 and 52.23 - Services incidental to transportation; and I56 - Food and beverage serving services. New indices are expected to be progressed for H52 and I56, improving coverage. Further work is needed in relation to G45 where international standards are yet to be developed.

The work team for the PPI is experienced and has low turnover. As far as possible, staff working on product groups that are included in both the CPI and PPI work with both indices. There are also regular quarterly result meetings with the National Accounts and the PPI team also participates in quarterly analysis sessions held by the Statistics Sweden Expert Group in economic statistics.

One area of weakness relates to the data processing system, where IT support is inexperienced in the system, and the potential for inadvertent error arises. One such error has been caught late in production.

Other accomplishments

In addition to the statistical developments in the service sector described above, the PPI team has continued to develop their Index Auditing Model to provide a structured way to view the quality of the index in total and for different sub groups. A score is generated, based on a number of criteria in each of five dimensions: sampling/coverage; pricing methods; specifications; response data; and quality adjustments. The aim is to

use as much quantitative data as possible, and maximise transparency. The audit model is particularly useful as an editing tool for lower level indices, and for prioritising development work.

The value of the Index Auditing Model will depend heavily on the quality of inputs in relation to each of the five dimensions and we encourage further development of these inputs, particularly objective measures.

Progress towards prior recommendations:

In summary, good progress has been made against most of the four recommendations from the previous report. In the case of the second recommendation, to monitor quality adjustments, an attempt was made but discontinued because of inadequate data, and a new recommendation is made in this round to pursue the appropriate data to implement this recommendation.

1. *Measure the Price of Trade Margins.* We asked the team to continue to develop the measurement of the price of trade margins. Future work has been planned using the Eurostat Grant proposal.

   This work has been taken forward and enabled Statistics Sweden to increase its coverage of the service sector with new indices to be progressed for warehousing and support services for transportation (H52) and Food and beverage serving services (I 56). A detailed research paper was also prepared for Trade and repair services of motor vehicles and motorcycles (G45) and an appropriate approach to progress this is still being considered independently of the Eurostat grant.

2. *Monitor Quality Adjustments*. We continued to support a recommendation for a measure comparable to the Implicit Quality Index of the CPI for the PPI.

   An attempt was made to develop an Implicit Quality Index for the PPI but was aborted because the approach taken did not sufficiently distinguish adjustments to prices based specifically on quality, from those adjustments made to account for other characteristics such as a change in volume.

3. *Coverage*. The team was asked to continue work to improve the coverage on the imports and exports of services as well as to investigate how to cover the new innovative products that enter the markets. We also asked that possible ways of updating the sample in relation to new businesses, services and new market products be investigated, along with plans to increase response rates for those groups where non-response is a problem for the National Accounts.

   Improved coverage of the PPI has been the focus of the work described under recommendation 1 above. It is noted that while the sampling frame is 2 years old, the products produced by these continuing businesses are observed in the survey and updated continuously. New products from businesses new in the last 2 years will not be covered. This may be an issue for areas of rapid product change and new production methods where relatively new firms may quickly build market share, for example as a result of government policy changes. The collection department has been working with strategies to limit non-response which is measured each period. A project analysing non-response to understand the breakdown of short term non-response (where imputation is more viable) and long term non-response, is planned for 2021. Resource constraints in 2020 led to reduced reminders by phone leading to some increase in non-response for the quarterly collection.

4. *Coherence and fitness for purpose.* We asked that the existing list of activities be developed into an action plan, to align the PPI and CPI methodologies and to increase the coherence between the two indices. We also asked that, coherence with other economic statistics important for macro-economic indicators, i.e. price statistics and the National Accounts, be explored.

PPI staff participate in a variety of activities and groups with the aim of ensuring coherence and fitness for purpose of the PPI. For product groups that appear in the CPI and PPI, staff work with both indices as far as possible. The unit shapes and takes part in the Voorburg Group Task Force – Framework for comparing SPPI and CPI methodologies. PPI representatives are involved in quarterly results meetings of the National Accounts and the quarterly analysis session organised by the Expert Group in economic statistics.

Key recommendations for the coming two years

1. *Index Auditing Model:* This is a very good initiative which provides a structured approach to understanding key sources of uncertainty in the index at various levels. It is being used as an output editing tool to raise issues of concern to senior staff, for training less experienced staff and for planning and prioritising improvement activity. The value of the index will depend heavily on the metadata input to it. It is recommended that there be continued development of these inputs for example the implicit quality indices, and measures of sampling uncertainty discussed separately below. It is also recommended that the audit model results be shared with users in discussions about quality requirements and achieved quality.

2. *Implicit quality indices:* An attempt to develop implicit quality indices was aborted as the approach being taken did not adequately distinguish those adjustments made as a result of perceived quality changes and those adjustments made as a result of a changed specification. It is recommended that the feasibility of calculating implicit quality indices that focus more purely on quality adjustments, be studied, with a view to better understanding where and by how much, quality adjustments are impacting data, and as an input to the quality audit index.

3. *Sampling uncertainty:* Confidence Intervals for the PPI were calculated in 2008, based on the sampling uncertainty. It is recommended that this exercise be repeated regularly at those levels of aggregation at which indices are provided to users. While the sampling uncertainty is not the only source of possible inaccuracy in the indices, it becomes a more significant source at lower levels of aggregation, and would be a useful and objective input to the Index Auditing Model for these aggregates.

4. *IT systems support:* There has been a full turnover of IT support staff for PPI, resulting in the potential for errors in the PPI. A joint review by the product area and IT of how to best mitigate this risk for ongoing production should be undertaken, and a strategy agreed.

5. *Documentation:* The Quality Report and the Production of the Statistics document, are clearly written and useful documents that allow users to understand key attributes of the PPI. It is recommended that significant developments such as the move from geometric means to arithmetic means at the lowest level of aggregation, and the backward evaluation of the impact of this, should be referenced in these documents to support users in fully understanding the product.

6. *Coverage of service statistics*: The PPI team is encouraged to continue its development work to increase coverage of service statistics both in product groups covered by the suite of PPI and XPI based outputs and also in the measurement of prices for international trade in services.

Other areas for consideration

Currently the two price areas, PPI and CPI operate as one unit, leading to substantial opportunity to ensure the coherence in methods, sharing of expertise and so forth between the units. Under the new organisation, the units will be formed into separate sections, although still within the same Economic Statistics and Analysis Department. Particular effort should be made to ensure the benefits of past team work, sharing of expertise and mechanisms to ensure coherence between the two sets of indices, are not lost in the new organisational arrangements.

Figure 2. PPI Ratings, Round 10

Sub and sub-subcomponents of Accuracy	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)
Overall accuracy	53	55	○	◒	○	◒	◓	◓	
Sources of uncertainty:	55	56							
-Sampling	55	55	◒	◒	◒	◓	◓	◓	H
-Frame coverage	58	60	◒	◒	○	○	○	○	M
-Measurement	53	58	◒	◒	○	○	◓	◓	H
-Non-response	57	52	◒	○	○	○	◓	◓	L
-Data processing	58	55	○	◒	◒	○	◓	◓	M
-Model assumptions	52	53	◒	◒	○	○	◓	◓	H
Preliminary statistics compared with final statistics	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	

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

3.3.3. Statistical Business Register, SBR

Context and discussion

The Statistical Business Register is Statistics Sweden's base register for businesses. It contains business units and linking variables to other base registers, as well as links to other statistical registers. The situation register is produced quarterly in March, May, August and November as a basis for frames for statistical surveys, and is the object of the ASPIRE review. The SBR now supports the following unit types: Legal Unit, Local Unit, Enterprise Unit, Kind of Activity Unit, and Local Kind of Activity Unit.

The variables in the register vary depending on the type of unit but the main types are: identification, demography, stratification information, links to other registers and relationships between unit types.
The register is updated with new legal entities as they are fed through from the Tax Agency. All units with more than one location are surveyed to update the population of local units. Profiling is the process for converting legal entities to Enterprise units and Kind of Activity Units. Profiling for large and complex units is done by the Large Case Unit and involves contact with the business. This profiling is regularly refreshed with continuous updating applied to the register as appropriate. Desktop profiling is undertaken for units below this cut off but with significant secondary or ancillary activity. Smaller units are assumed to be have a simple structure.

Relationships between the Business Register Unit and its key providers, particularly the Tax Agency, are good. The flow of information from the Tax Agency to the SBR data is timely and there has been no impact of COVID-19 to the smooth flow of data.

In addition to accurately reflecting the number and structure of businesses operating in Sweden, key quality issues for the SBR in terms of its use as a frame for business surveys are the accuracy and completeness of auxiliary information, particularly industry codes (provided by the Tax Agency), size information (turnover from VAT, and employment by location from PAYE), and activity status (i.e. whether economically active.) While these are all believed to be of good quality and are amended as appropriate through maintenance procedures, there do not appear to be structured processes for feeding back information from survey areas where this can be used in the register, for example for units from the take all strata of surveys. Nor is there regular feedback from survey areas of estimated quality of industry codes, size, or activity status flags.

There is regular information provided by the SBR to business survey units with each quarterly snap shot of the frame, on the changes in population demographics. However, areas conducting sub-annual surveys, do not use quarterly updates of their frames even though these are available. Instead, they update their samples annually.

Other accomplishments

Since the last ASPIRE round, there have been three key developments for SBR. The first is the re-engineering of the SBR processing system. The new BR database has been launched together with a new application for loading and processing data, making loading more transparent and flexible, and better able to meet user needs. Data loading has sped up processing and allows increased focus on organisational change in businesses, as well as prioritising of units that are most important, and allowing for more senior management input.

The second is the addition of the monthly PAYE tax return data to the register providing a timely employment related size measure and a source of employment data where employment number is missing from the register. Changes in the monthly data, by location have the potential to be a quick indicator of units that have changed structure for maintenance follow up.

Thirdly, profiling of the largest, most complex business units (about 180 enterprise groups) by direct contact, has continued and a major project, SAMSTAT has been set up to implement the changed structures in business surveys, with the aim of all statistics with 2022 reference year being based on the profiled population. Units with large secondary or ancillary activities will be desk profiled (about 470 enterprise groups and 300 municipalities) and for the remainder, a simple structure will be assumed with group equalling enterprise.

Progress towards prior recommendations

1. *Conclude the re-engineering project and the profiling of large businesses. Develop an implementation strategy for profiling*: Re-engineering was completed Nov 2019. Profiling has continued and the SAMSTAT project has developed an implementation strategy for business surveys. Work on backcasting data to 2000 based on the Kind of Activity Unit of profiled units is underway. The SBS (enterprise based) is already using profiled units.

2. *We asked the SBR unit to liaise with survey areas on key requirements to ensure quarterly frames are usable and used. This would include implementation of an activity status code, and a centralised approach to frame validation to reconcile differences between quarterly frames. We also asked SBR to set up quality monitors of the SBR from those surveys using it.* Although the implementation of PAYE data does provide a further source for setting an activity status code, to date other priorities have taken precedence over both components of this recommendation. Feedback of register related information from business survey areas to the SBR is limited.

3. *A review of the size and design of the business units survey in the light of the potential for automating detection of likely new locations using employee address information was recommended*. Units are added to the business units survey if auxiliary information indicates a unit recorded with one location may actually have multiple locations. PAYE will be a fast source of this information from 2020 on.

Key recommendations for the coming two years

1. A strategy for evaluating the impact of profiling large businesses on the quality of business surveys should be considered with a view to judging the effort appropriate to allocate for the ongoing profiling and associated maintenance program.

2. The business register unit should work with business areas running sub annual surveys to identify the impediments to using the quarterly refresh of their frames, with a view to reducing these impediments.

3. The business register unit should develop structured processes for the feedback to the register of information obtained from surveys that is able to be fed back. In particular feedback relating to large businesses falling within the 'Take all' strata of surveys should be fed back to the register where the quality of the information is judged adequate.

4. The feasibility of collecting quality monitoring information for the SBR from business survey sources should be investigated. This would include measures such as percent misclassified to industry by size, and correct industry; percent with missing or misleading size information; and percent inactive by industry and size. This would help ensure the quality of the register information remained 'in control' and help prioritise improvement activities.

5. The business register has made some progress on a web application to reduce respondent load in reporting changes to business location information. This initiative should be pursued in line with corporate infrastructure development.

Other areas for consideration

The business area should continue to be proactive in helping build the relationship between their data providers, particularly the Tax Agency, and Statistics Sweden.

Figure 3. SBR Ratings, Round 10

Sub and sub-subcomponents of Accuracy	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)
Overall accuracy	55	63	◒	◒	○	◒	○	○	
Sources of uncertainty:	58	61							
-Sampling	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
-Coverage	62	62	◒	◒	○	○	◓	○	H
-Measurement	58	62	◒	◒	○	○	○	○	H
-Non-response	47	40	○	○	◓	◓	●	○	L
-Data processing	62	65	◒	○	◒	◒	○	◒	H
-Model assumptions	52	63	◒	◒	○	◒	○	○	M
Preliminary register compared with final register	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	

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

3.3.4. Quarterly Gross Domestic Product, GDP-Q

Context and discussion

The GDP-Q estimates are produced from a very large and complex set of inputs and are compiled using recognised international standards. They provide the fundamental measures of growth for the economy and are among Statistics Sweden's key products.

In this review, we looked particularly at the impact of COVID-19 on the quality of GDP estimates and the use of turnover data to model household consumption.

The SAMSTAT project, which is introducing profiling of large companies in the Statistical Business Register, will produce major benefits for national accounts who are a major customer for this work.

The period since our last review has been one of significant development for GDP-Q, which we summarise below. Earlier publication of the estimates will, in particular, have been welcomed by users, and have enhanced Statistics Sweden's reputation. The recommendations that we make are focused on some fundamental aspects of the accounts and will require close working with other areas of the office, including business surveys, and IT. The new organisation structure should help this.

Progress towards prior recommendations

1. *Household Consumption Data*. In the absence of a Household Budget Survey a sufficient and more robust data source solution needed to be found, using the experience of other statistical offices. We suggested this might come from the supply side and sources such as scanner data should be investigated. A Task Force has been formed to take forward the development of estimates based on turnover data.
2. *Quarterly Business Survey*. We were concerned that in the absence of a survey, quarterly GDP may not accurately pick up turning points in the economy. We recommended that consideration be given to carrying out this survey as soon as possible. We understand that work has progressed on the design of a survey which will begin next year. In the meantime, estimates are being made using VAT data.
3. *Implementation strategy for the new IT system*. We were concerned that implementing the new IT system could add to risk, and national accounts needed to start thinking about how best to mitigate against the risk to their operations. We have learned that implementation of the new IT system has been delayed. This year compilations are being made in both the old and new systems, to ensure data quality.

Other accomplishments

1. A 'flash estimate', an early estimate of the quarterly estimate, together with monthly estimates, began to be published from May 2020, advancing publication by three weeks from the regular publication. This has driven improvements in timeliness of some source data, e.g. local authority data.

2. Adjustments have been made to components of the national accounts, for example output for the health sector, to allow for changes in economic behaviour due to the pandemic.

3. Work has taken place to harmonise the National Accounts with the Balance of Payments, which has improved the quality of both sets of statistics. Work is continuing to harmonise with the financial accounts.

4. Quarterly data from the VAT system is being used as a source for a number of industries and to validate production and value added for the business sector, improving the quality of the flash estimates and the quarterly estimates.

5. Co-operation with the Large Case Unit associated with the SBR has improved the coherence and the quality of some of the source data.

6. Quarterly meetings of all producers of economic statistics are now held to improve the coherence of the economic picture.

7. There has been a significant advance in the publication of the annual accounts, by six months, which is in time to inform the autumn budget. Publication dates for quarterly GDP have been co-ordinated with the Production Value Index and the Household consumption indicator, enabling a coherent story to be told of the drivers of economic growth.

Key recommendations for the coming two years

1. *COVID-19*
   a. The pandemic is thought to have had only a slight effect on the quality of the national accounts, with response rates of business surveys holding up well, and most of the impact being seen in changes to consumer and business behaviour, which require changes to modelling assumptions. These will need to be reviewed carefully as the economy recovers to ensure they remain relevant.

   b. There will be a need to work with survey sources to use business demography data on births to ensure these are picked up as the economy recovers.

   c. There will be a need to manage the almost inevitable increase in volatility of seasonally adjusted data that the disruption to seasonality and trading day effects etc. will cause, probably for a number of quarters to come. There is a need to ensure methodological resources are adequate to explore impacts and alternative estimation strategies.

   d. There is a need for the team to be open in communication with users about these issues.

2. *Quarterly Structural Business Survey*. We note the progress being made to introduce a new survey next year as part of the redesign of Economic Statistics. Given the concern that in its absence, quarterly GDP may not accurately pick up turning points in the economy, we recommend that priority continue to be given to this work.

3. *Household consumption*. Validate the new method using turnover data to ensure that it is a good approximation to household consumption.

4. *New IT systems*. Carefully manage the delivery of the new system, ensuring adequate acceptance testing of the system and the documentation for those left to support it, and training of the GDP team. Ensure that national accounts has access to continuing IT expertise for post implementation modifications, and monitor its impact on the statistics.

5. *Regular revisions analysis*. Use the new IT system to produce regular analysis for users of revisions, by size and type, over time.

6. *Quality report: non-response*. State explicitly in the Quality Report for GDP that non-response in surveys has an impact on GDP-Q and is discussed in the QDs of surveys; and state explicitly the initiatives that are being tackled collaboratively to mitigate these impacts on GDP-Q.

Figure 4. GDP-Q Ratings, Round 10

Sub and sub-subcomponents of Accuracy	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)
Overall accuracy	55	58	○	○	○	◒	◓	○	
Sources of uncertainty:	55	57							
-Sampling	42	47	◒	◓	○	○	●	○	L
-Frame coverage	62	52	◒	◓	○	◒	●	○	M
-Measurement	52	65	◒	◒	○	◒	○	○	H
-Non-response	47	47	○	○	○	◓	◓	○	L
-Data processing	55	55	○	◒	○	◒	◓	○	M
-Model assumptions	62	62	◒	○	○	◒	○	○	H
Preliminary statistics compared with final statistics	67	63	◒	○	◒	◒	○	○	

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

4. Cross-Cutting Issues and Recommendations

4.1 Statistical Leadership

The new organisational structure provides the stimulus to develop statistical leadership throughout the organisation. We encourage Statistics Sweden to:

1. Use the opportunity of the reorganisation to encourage statistical leadership at all levels of the organisation, so that each subject matter area is outward looking, able to spend time with users, identify new opportunities for relevance, and problem solve issues. Use the central methodological unit to actively build the methodological capability across the agency and to ensure a methodological presence in corporate decision making.
2. Ensure there is clear messaging to staff, especially senior staff, about the role, goals and values of the organisation so that they are empowered to make decisions and take risks in a dynamic environment in line with the roles, goals and values, and, where possible back this up with clear policy statements.
3. Ensure those involved in 'business as usual' understand where the organisation is going and why, and are kept up to date about progress being made. Risks associated with significant products that will continue to operate on tight budgets and legacy systems while new approaches and infrastructure are developed, will need to be actively managed.

4.2 The impact of COVID-19

The COVID-19 pandemic has had a significant impact on statistical production. We encourage the office to:

1. Reflect on how this experience can help them adapt to the vision of a more responsive agency they are now pursuing.
2. Ensure that there is very good communication with users on resulting quality issues around the statistics, so that, in the event of future revisions, users are aware of some of the difficulties in making good estimates at this time, and feel they have had a chance to input to solutions where relevant.

4.3 Common processes and tools to support a consistent and current view of the business population across all surveys

The SBR provides a very current view of the business population and its key attributes at any time, however it seems that survey areas each use this to develop a particular view that requires resources to validate and maintain with each frame update. This resource requirement appears to prohibit quarterly updates to frames for sub-annual surveys. We encourage Statistics Sweden to take an 'across survey' view of whether tailored solutions for each survey are warranted, in terms of the statistical benefits and costs as well as the resource costs.

4.4 IT Constraints

Moving towards common tools and systems will in the longer term, lead to more effective solutions with reduced overall maintenance costs. This will help alleviate the current problem of competing demands and reduce the risk to quality posed by a shortage of IT resources.

The first step in building common technical infrastructure will be to analyse business processes and harmonise surveys and collections. This will avoid the high cost of tailoring each common tool to very specific processes. The failure of the individual business surveys to update frames quarterly despite the corporate tool being available (see 4.3 above), is an example of this issue. However even with harmonised surveys and collections, there is a need to provide functionality that allows each product area to work effectively with the corporate tools.

4.5 Development of data science and machine learning skills

The data revolution has sparked a growing need for data science skills and machine learning, and we encourage Statistics Sweden to further develop expertise in data science and machine learning.

4.6 Household expenditure data

Household expenditure data is a key source for consumers' expenditure (a key component of GDP), and for the high level weights for the CPI. Although alternative sources have been found and work is progressing, none of these sources directly measure what consumers spend.

We encourage Statistics Sweden to consider what options there are to collect spending data from households that could reach an acceptable level of quality.

5. References

Linacre, Susan, and Stephen Penneck. 2020. *A Ninth Application of the ASPIRE Quality Evaluation System for Statistics Sweden*. Accessed June 15, 2021. A Ninth Application of ASPIRE at Statistics Sweden.pdf (scb.se)

Office for National Statistics. 2021. *Coronavirus and the latest indicators for the UK economy and society methodology*. Accessed June 15, 2021. Coronavirus and the latest indicators for the UK economy and society methodology.

Office for Statistics Regulation. 2021. *Statistical leadership: Making analytical insight count.* Accessed June 15, 2021. Statistical leadership: Making analytical insight count.

Reinsdorf, Marshall. 2020. IMF Working Paper. *COVID-19 and the CPI: Is Inflation Underestimated?* Accessed June 15, 2021. COVID-19 and the CPI: Is Inflation Underestimated?

Statistics Sweden. 2021. A System for Product Improvement, Review and Evaluation, ASPIRE – a manual. Accessed June 15, 2021. A System for Product Improvement, Review and Evaluation, ASPIRE - a manual (scb.se)