Quality Indicators Poster Session

Index Auditing Model – Swedish experience

Authors

Rohan Draper

Marcus Fridén

Statistics Sweden

Statistiska centralbyrån

Contents

1. Background
- – Overview
- – Purpose of the Model
- – Collaboration: Statistics Denmark

2. Model Description
- – Valuation Criteria
- – Categories

3. Demo - Voorburg 2017
- – Excel version

4. What's next…
- – Q&A Session

Key Outcomes

- • Update on the development work and application of Sweden's Index Auditing Model.
- • Demonstration of the model

1. Background
Overview

- Statistics Sweden presented at the Nordic Seminar 2016 a prototype index auditing model and an applied version at the Nordic Seminar 2017.

- Practical application on the Swedish PPIs (incl. SPPIs) throughout 2016/17.

- International collaboration with Statistics Denmark as a reference country (three formal workshops) and model comparison with Statistics Canada (one formal meeting).

1. Background
Purpose of the Model

- Increase the understanding of price indices (and their quality) amongst users

- Increase our own understanding of the types of process data that we want to collect and analyse

- Give input to our internal users auditing (for example, National Accounts)

- Understand where overviews are likely needed (evidence based)

- Provide an transparent consistent basis for resource allocation / planning

- Encourage discussion about important end-to-end aspects of index production and output

1. Background
Collaboration – Statistics Denmark as a reference country

Three meetings –

1) Discussion of auditing metrics and a review of the auditing methodologies.

2) Joint application (as a benchmarking activity); discussion of findings; and comparison of results.

3) Consultative review of documentation and learnings from practical application.

Learnings from implementation –

- Challenges and opportunities
- Potential application

2. Model Description
Valuation Criteria

1 = Low quality / high risk; the index likely has a clear bias, index review required

2 = Low-medium quality / medium risk; index should be representative in the long run but may be misleading in some periods; index review required

3 = Ok, approved (not prioritised for immediate review)

4 = Good quality / low risk. High confidence in the representativeness of this index; not a review priority

5 = Excellent quality / low risk. Not a review priority.

The aim of the model is to use as much quantitative information as possible. This is to maximise transparency, repeatability and comparability of the audit. Out of necessity some categories include both subjective and quantitative valuation methods. Valuations should be motivated with descriptive justifications.

2. Model Description
Categories

1. Sampling/ Coverage
2. Pricing Methods
3. Specifications
4. Response Data
5. Quality Adjustments

Categories were chosen with price statistics directly considered. Other standardised categories were researched from existing quality models, however, it was determined that these standard models did not fit the practical needs of price statistics.

The categories broadly cover three stages of index production: pre provider contact; provider contact (initialisation); and post provider contact (review of reported data).

3. Demo – Voorburg 2017

- The following demo is based on the version of the model dated 2017-09-01 and presented as a poster session.

Index Auditing Model

Background

Our aim is to capture a representative, average, pure price movement (adjusted for volume and quality changes) for all product groups (or industries) and weighted to the total for the entire economy in each measurement period. A current gap in general price index theory and practise is quantifying in a meaningful way how well we are meeting this goal (or the risk that we are not). This model has a general purpose to provide an answer to this challenge with a focus on the quality of individual indexes. A score is generated based on a number of criteria applied to a x-digit level index. Results are then reviewed both at the individual index level and at the aggregate level.

Purpose

Primary:
- Understand where index reviews are likely needed (evidence based)
- Provide a transparent consistent basis for resource allocation / planning
- Give input to our internal users' own auditing (for example, National Accounts)

Secondary
- Increase the understanding of price indices (and their quality) amongst users
- Increase our own understanding of the types of process data that are able to be collected and analysed
- Encourage discussion about important end-to-end aspects of index production and output

Categories
1. Sampling/Coverage
2. Pricing Methods
3. Specifications
4. Response Data
5. Quality Adjustments

Valuation Grade

1 = Low quality / high risk; the index likely has a clear bias, index review required

2 = Low-medium quality / medium risk; index should be representative in the long term but may be misleading in some periods; index review required

3 = Ok, approved (not prioritised for immediate review)

4 = Good quality / low risk. High confidence in the representativeness of this index; not a review priority

5 = Excellent quality / low risk. Not a review priority.

The aim of the model is to use as much quantitative information as possible. This is to maximise transparency, repeatability and comparability of the audit. Out of necessity some categories include both subjective and quantitative valuation methods. Valuations should be motivated with descriptive justifications.

Output

Index review at the x-digit level of aggregation with detailed assessment (2-digit level in Sweden)

Audit summary enabling assessment of all index results (results by index / results by criteria)

Visualiation / compilation that enables comparison of different assessment periods to review progress

1. Sampling/Coverage

Valuation Criteria	Description of current status	Reference Data	Guidance	Grading Motivation	Grade
Sampling Method	PPI General - PPS		Does the selection method allow all objects in the population to be drawn? Is there a risk that we will miss a significant part of the price trend otherwise? Is it possible to calculate		5
Sampling Frequency	PPI General - Annually		The higher the sampling frequency the higher the likelihood that the index is representative of actual market activity which results in a higher rating.		5
Sample Frame's Timeliness	PPI General - Frame is two years old (y-2) SPPI General - Frame is three years old (y-3)		The older the frame is that greater the risk that the frame isn't representative of reality. An older frame is equivalent to a low rating.		4
Attrition		Process data from annual update work (code 7).	A high attrition rate at the sample selection process leads to uncertainty that we are measuring accurately what we want to be measuring. A high attrition rate equates to a low rating.		3
Coverage		Proportion of total production being directly covered.	The higher the coverage the higher certainty we have that we accurately represent market activity. A higher coverage equates to a high rating.		3
Random sampling error		Variance/ standard deviation calculated from process data from annual update work.	Higher variance is equivalent to higher uncertainty and is considered as a risk giving a lower result.		3
Cut-off		Proportion of the population outside of the cut-off	The higher the proportion of the population that is outside the cut-off the lower the certainty we have that we accurately represent market activity.		3
Frame error		Process data from annual update work (code 8,9 and 8X).	A high percentage frame error results in higher uncertainty and a lower rating.		3
Valuation Grade	*<How reliable do we estimate that the development of the index has been using the above guidance?>*		Comments:	*<comments included here will be included in the summary report>*	3,63

2. Pricing Methods

Valuation Criteria	Description of current status	Reference Data	Guidance	Grading Motivation	Grade
Actual prices - "direct use of prices of repeated products" (excluding contract prices)		Proportion actual prices.	A high percentage equates to a high grade. That is, a high usage of actual transaction prices of repeated products equates to low risk.		5
Contract Prices		Proportion contract prices.	3-6 contracts with varying renewal periods (per reporting provider) are required for high ratings. A low number of contracts with non-varied renewals results in a low rating.		3
Transfer prices		Proportion transfer prices.	The rating is user dependent. For NA, the use of transfer prices is approved. On the other hand, if it is for contractual regulation, it is not approved.		3
Estimated prices using related observed prices (Component and/or percentage fee prices)		Proportion related observed prices.	High percentage of estimated prices as well as specifications that are not regularly updated give low ratings. Quality issues with underlying index also gives a low rating.		3
Hypothetical prices (Model prices)		Proportion model prices.	High percentage of model prices and models that are not regularly updated gives a low rating.		3
List prices		Proportion list prices.	A high percentage of list prices results in a low rating.		3
Time based prices		Proportion time based prices.	A high percentage of time based prices results in a low rating.		4
Unit Value prices		Proportion unit value based prices.	An average price is harmful if the products included are not sufficiently homogeneous.		3
Valuation Grade	*<How reliable do we estimate that the development of the index has been using the above guidance?>*		Comments:	*<comments included here will be included in the summary report>*	3,38

3. Specifications

| Valuation Criteria	Description of current status	Reference Data	Guidance	Grading Motivation	Grade
How often can specification updates be actioned	PPI General: We have full flexibility to update/ change specifications from month to month. A restriction is the possibility to increase the number of specifications from a particular business between annual updates.				4
Are specifications too broadly defined? (i.e. risk of product mix / undetected quality change)		1. Price volatility calculation 2. Quality check of specifications (manual check)	Specifications that are *too broadly defined* are regarded as not having sufficient detail to be able to adequately observe quality and may result in false price changes impacting results.		4
Are specifications too tightly defined? (i.e. low transaction frequency / high product change frequency)		1. % - product change (21:or) 2. % - non reporting / no transaction (55:or)	Specifications that are too tightly defined may have overly detailed parameters that make it difficult for companies to find comparable transactions leading to a higher frequency of product changes, non reporting and/or absence of transactions.		4
Valuation Grade	*<How reliable do we estimate that the development of the index has been using the above guidance?>*		Comments:	*<comments included here will be included in the summary report>*	**4,00**

4. Response Data

| Valuation Criteria	Description of current status	Reference Data	Guidance	Grading Motivation	Grade
Imputation Methods	PPI General - default method is imputation of an average movement from the closest aggregate with more than three reported specifications. Manual imputation is also utilised.				5
Inliers. No price change.		% of specifications that have reported no price change for X number of periods (code 11)	Inliers refers to specifications that have falsely showed no price change over time. A high level of no price changes is an indicator of possible error.		4
Unobserved data and non-response	1. 2. 3. 4.	1. % - 55 (no transaction) 2. % - 56 (non-response) 3. % - 57 (discontinued) 4. Total % - 55,56,57.	A large amount of unobserved data leads to uncertainty and potential bias. A high percentage results in a lower rating.		5
Valuation Grade	*<How reliable do we estimate that the development of the index has been using the above guidance?>*		Comments:	*<comments included here will be included in the summary report>*	**4,67**

5. Quality Adjustments

Valuation Criteria	Description of current status	Reference Data	Guidance	Grading Motivation	Grade
Proportion of specifications updated (product change / product update)		Proportion of specifications updated	Updating specifications is important to maintain representativeness. However, changes can add subjectivity to price measurements. A high % of changes is considered a potential risk.		3
Quality Adjustment Method	1. 2.	1. % - implicit 2. % - explicit	Ideally quality adjustments are well informed. This criteria sheds light on the prevalence of implicit and explicit methods. A high % of for example carry forward adjustments is deemed a risk.		4
Weighted effect of quality change		Proportion of index develop for X reporting periods that is attributable to quality change.	An overly large influence from quality change or no prevalence of quality change at all are both indicators that		2
Quality adjustment method (1)**					
Quality adjustment method (2)**					
Quality adjustment method (3)**					
Quality adjustment method (4)**					
Quality adjustment method (5)**					
Valuation Grade	*<How reliable do we estimate that the development of the index has been using the above guidance?>* ***specific quality adjustment methods can be added on a needs basis. Otherwise the summary version of the question is asked delineating usage of explicit versus implicit methods.*	Comments:	*<comments included here will be included in the summary report>*		3,00

Index Auditing - SPIN XX
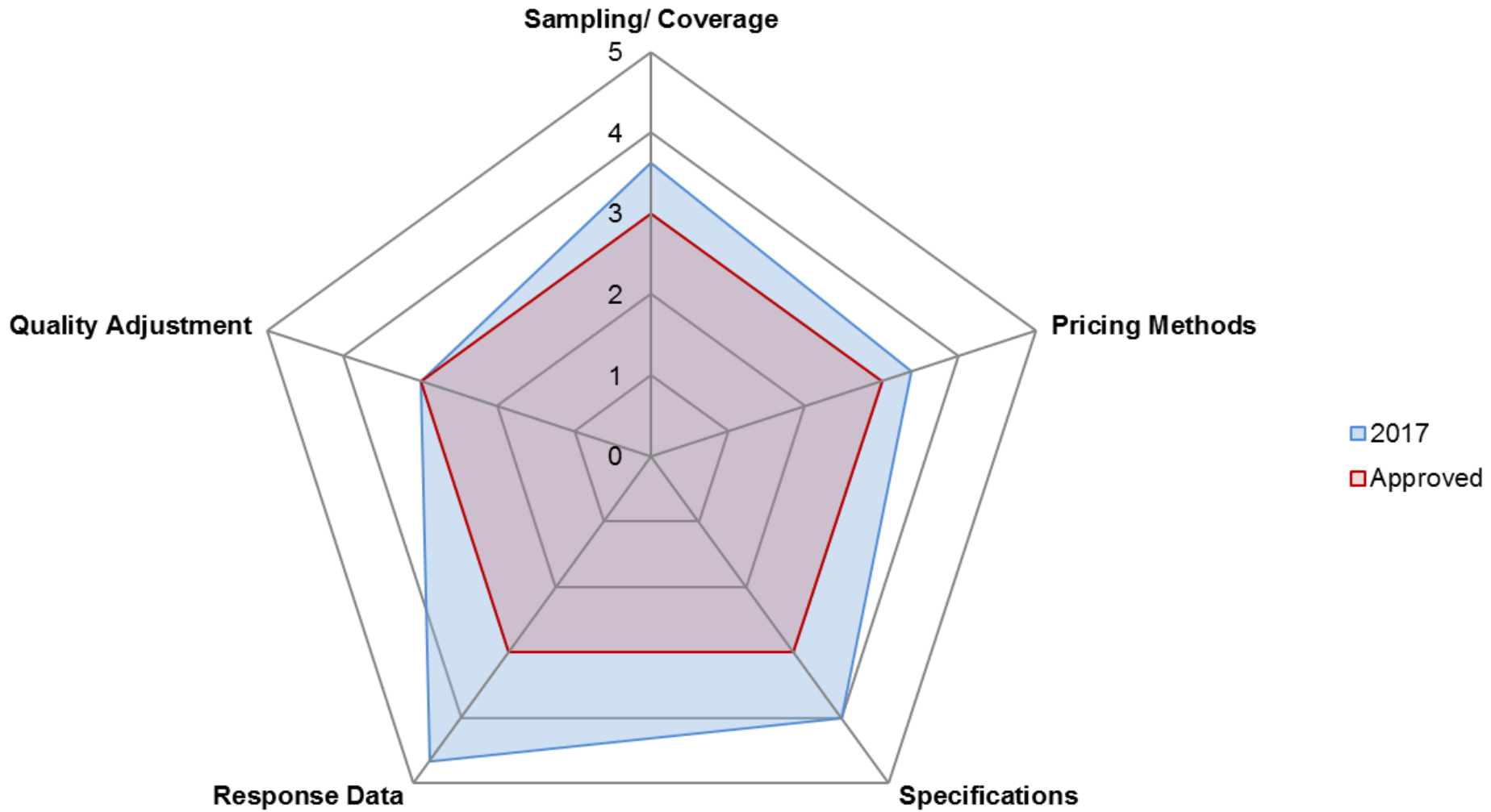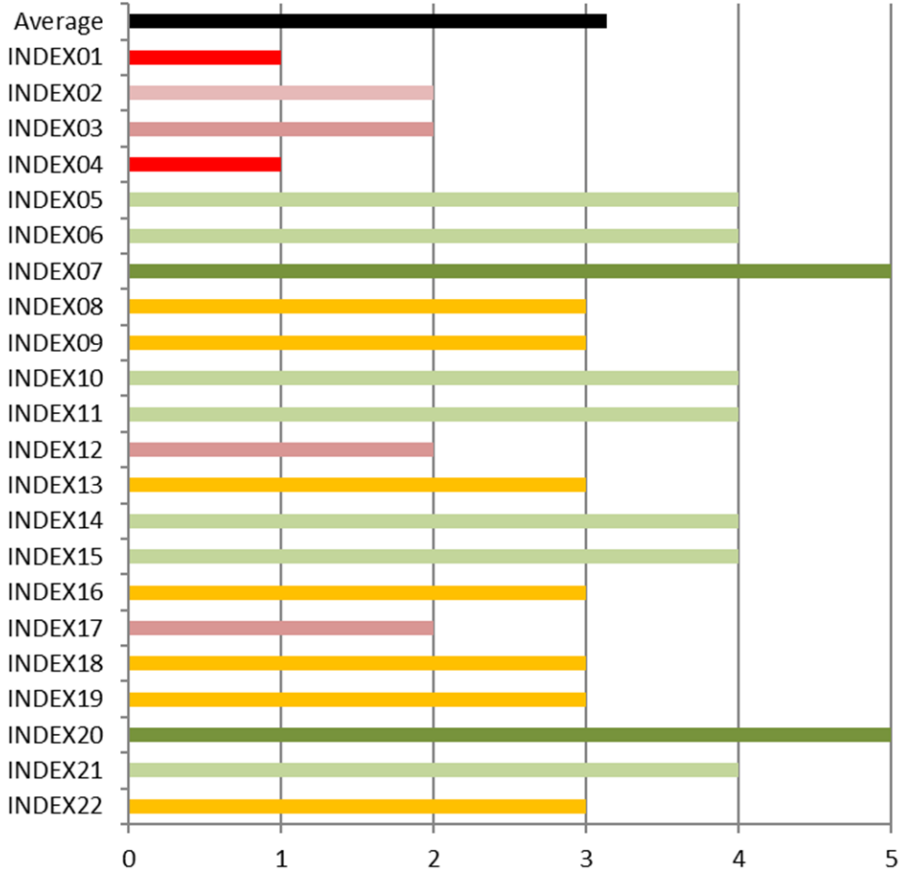
Challenges –

- Development of automated metadata that supports index audits. *[Statistics Sweden uses a purpose built compilation software and this software is flexible to add-ons like SAS programmes that enables inclusion and extraction of metadata]*

- Is there still an element of untoward subjectivity? Eliminating subjectivity requires appropriate staff training combined with process data and auditing tools. *[At Statistics Sweden we have developed a prices course plus a self-assessment tool that compliments our index auditing activities]*.

Application –

- Internal application. Statistics Sweden intends to incorporate index auditing as part of our annual statistical quality practices and initial results have already been presented to key users.

- Is there a use for country to country comparison? If so standard "guidance" and rules will be required.

Statistics Sweden

Statistiska centralbyrån