WORKING PAPER No. 21
February 1993

THE INDEX OF INDUSTRIAL PRODUCTION
A Formal Description of the Process Behind It

Claes-Håkan Gustafsson
Åke Holmén

Svensk resumé på sid 34.

THE INDEX OF INDUSTRIAL PRODUCTION
A Formal Description of the Process Behind It

CLAES-HÅKAN GUSTAFSON and ÅKE HOLMÉN

ABSTRACT

The index of industrial production, IPI, is published in several different versions from the first preliminary version over a number of revisions until the final definite version, which has a time lag of more than two years. This paper presents an overview of all these versions. It provides a formal description of the methods of the different revisions with respect to the input data used as well as to the computational methods. The paper also provides a check of whether the process involved in the production of the monthly IPI is under control in a statistical sense. It turns out that this is not at all obvious, at least not on a disaggregate level.

1. INTRODUCTION

Forecasts and analyses of various economic phenomena are based on input data of different kinds. Regardless if they are generated by models or have more informal elements, their quality depends on the quality of input data. Stabilisation policy is also based on the data available at the time of the decision. Therefore, the quality of the data obviously has an effect on the degree of success of the chosen policy.

Our paper deals with the index of industrial production (IPI) produced by Statistics Sweden. The IPI is a volume index designed to measure the development of industrial production in the Swedish industry[1] on a monthly, quarterly, and yearly basis.

The IPI is published in a number of versions from the initial preliminary version to the final definite one. A user of IPI about to make a forecast may feel tempted to take the latest preliminary version available. The revisions of IPI that follow may then necessitate

[1] I.e. branches 2 and 3 in the SNI–code.

alterations in the forecast and subsequent changes in decisions. These revisions may in other words constitute a quality problem, which justifies this documentation and critical analysis of the methods used in deriving the different versions of the index.

Our aim, therefore, has been to describe and formalise the methods used in producing different versions of monthly, quarterly, and annual IPI.

2. DATA

The preliminary monthly index is computed from sample based production indicators, which may be of three kinds (Statistical Report, Series I, July 1991, p 2), depending on what branch the index is computed for, namely

1. Production data
2. Employment data (number of hours worked)
3. Delivery data

These data are supplied by Statistics Sweden or by various branch organisations. Furthermore, input data from the latest annual statistics and from the quarterly delivery and inventory statistics are used in the derivation of the index. Successive revisions are made whenever more accurate input data become available. We now turn to a description of the sources of input data, time lags, revisions, and notations.

2.1. Data on production

Monthly data are based on samples and are supplied mainly by branch organisations. The time lag is about one and a half months. No revision is made of these data. We denote production of the i:th product, in month m year t, Q^m_{it} (i = 1, ..., n).

2.2. Data on employment

The employment data used concern the number of hours worked. These figures are based on samples and are supplied by the Department of Labour Statistics at Statistics Sweden. The preliminary version has a time lag of one and a half months. After another two weeks a revised (final) version is available. The preliminary version of the number of hours worked is denoted a^{m*}_t and the revised version a^m_t.

2.3. Data on deliveries

These data are sample estimates and are taken from the short-term statistics produced by Statistics Sweden. The method of collecting data makes several versions of the data available. Every month the firms in the sample report data on deliveries for the last three months. This means that data on deliveries for any given month will be available in three versions. The time lag of the preliminary version is about one and a half months. The first revised version is published in about two and a half months, at the same time as the preliminary version of the following month is available. Finally, the second revised (final) version of delivery data becomes available three and a half months after the end of the month concerned.

We use the following notations for the three versions of delivery data for month m year t: Preliminary version, l^{m**}_t, revised version, l^{m*}_t, and final version, l^{m}_t. Deliveries are measured in current prices.

2.4. Inventory data

The inventory data used as inputs are sample estimates and are collected by Statistics Sweden on a quarterly basis. The firms in the sample report the stocks of inventories of goods in process and finished goods both at the beginning and at the end of the period. Aggregate inventory investment is easily derived as the difference between stocks at the end and at the beginning of the period. The time lag of the first version is two months from the end of the quarter concerned. The final version is published three months later, i.e. two months after the end of the following quarter. The preliminary version of the aggregate inventory investment in quarter q year t is denoted $\Delta S^{*}_{t,q}$, and the definite version $\Delta S_{t,q}$.

2.5. Annual data

The annual industrial statistics are based on a survey of all industrial branches. Every plant in each branch reports the total production of that year of product i, denoted Q^{g}_{it}. Information about the market value of the output, $(PQ)^{g}_{it}$, is also collected. The annual statistics have a time lag of 15 – 18 months because of the extensive work required for collecting data, dealing with non-response, etc.

3. OVERVIEW OF THE REVISIONS

A preliminary monthly or quarterly index is revised a number of times, whenever revised input data become available for the time period concerned. From now on we concentrate on the regular revisions. Additional revisions are made with irregular intervals. If input data are revised outside the regular intervals described above, this may bring about extra revisions of the indexes. Another reason may be a change in the base year of the index.

3.1 The monthly index

A monthly index may be revised between three and six times, depending on what branch it concerns. The revisions caused by new quarterly or annual data are denoted "Q–revisions" and "Y–revisions" respectively. The following notations will be used:

A. Preliminary monthly index	$I^m_{A,0t}$
B. Q–revision 1	$I^m_{B,0t}$
C. Q–revision 2	$I^m_{C,0t}$
D. Y–revision 1	$I^m_{D,0t}$
E. Y–revision 2	$I^m_{E,0t}$
F. Final index	$I^m_{F,0t}$

Thus, $I^m_{B,0t}$ denotes the first Q–revision of the monthly index for month m year t with the base year 0. In branches where production data are used as indicators, there is no Q–revision of the monthly index. Therefore, versions B and C are not published for these branches. Y–revision 2 is made for months in the first half of the year only. Table 1 presents the schedule of revisions of all monthly indexes in year t.

If the times of publication of different versions coincide, only the most "up to date" version is published. This means, for instance, that a second Q–revision of the months in the first quarter will never be published, since data for the first Y–revision are available at that time. The different versions are, with a few exceptions, published chronologically from version A to F. The only exceptions are the months in the second quarter. For these months the first Y–revision is made in September. The final quarterly statistics of deliveries and inventories, which make the second Q–revision possible, are not available until December. Therefore, instead of waiting for these data the chronology of publication is interrupted.

Table 1. Months of publication of preliminary and revised monthly indexes of industrial production, IPI.

Month concerned	Prel. index ($I^m_{A,0t}$)	Q– rev.1 ($I^m_{B,0t}$)	Q– rev.2 ($I^m_{C,0t}$)	Y– rev.1 ($I^m_{D,0t}$)	Y– rev.2 ($I^m_{E,0t}$)	Final index ($I^m_{F,0t}$)
Jan_t	Mar_t	Jun_t	Sep_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Feb_t	Apr_t	Jun_t	Sep_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Mar_t	Maj_t	Jun_t	Sep_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Apr_t	Jun_t	Sep_t	Dec_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Maj_t	Jul_t	Sep_t	Dec_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Jun_t	Aug_t	Sep_t	Dec_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Jul_t	Sep_t	Dec_t	Mar_{t+1}	Sep_{t+1}	–	Sep_{t+2}
Aug_t	Okt_t	Dec_t	Mar_{t+1}	Sep_{t+1}	–	Sep_{t+2}
Sep_t	Nov_t	Dec_t	Mar_{t+1}	Sep_{t+1}	–	Sep_{t+2}
Okt_t	Dec_t	Mar_{t+1}	Jun_{t+1}	Sep_{t+1}	–	Sep_{t+2}
Nov_t	Jan_{t+1}	Mar_{t+1}	Jun_{t+1}	Sep_{t+1}	–	Sep_{t+2}
Dec_t	Feb_{t+1}	Mar_{t+1}	Jun_{t+1}	Sep_{t+1}	–	Sep_{t+2}

3.2. The quarterly index

Depending on the branch concerned, a quarterly index may be published in three to six versions denoted in the following way:

A. Prel. quarterly index (version 1)	$I^q_{A,0t}$
B. Quarterly index (version 2)	$I^q_{B,0t}$
C. Quarterly index (version 3)	$I^q_{C,0t}$
D. Y–revision 1	$I^q_{D,0t}$
E. Y–revision 2	$I^q_{E,0t}$
F. Final quarterly index	$I^q_{F,0t}$

For branches where production data are used as indicators, versions B and C of the preliminary quarterly index are not computed. Y–revision 2 is made for quarters in the first half of the year only. Table 2 shows the schedule of revisions and dates of publication of the quarterly indexes.

As was the case of monthly indexes, only the later version of the quarterly index is published, if the dates of publication for two separate versions happen to coincide. The publications are chronological, with the exception of the index for the second quarter.

Tables 1 and 2 give the dates of publication in Statistics Sweden's time series data base (TSDB) and, from 1992, in the monthly publication "Indicators". Older IPI–data are published in Statistical Report, Series I (SM:I).

Table 2. Months of publication of preliminary and revised quarterly indexes of industrial production, IPI.

Quarter concerned	Prel. vers.1 ($I^q_{A,0t}$)	Prel. vers.2 ($I^q_{B,0t}$)	Prel. vers.3 ($I^q_{C,0t}$)	Y-rev.1 ($I^q_{D,0t}$)	Y-rev.2 ($I^q_{E,0t}$)	Final index ($I^q_{F,0t}$)
Quarter 1	Maj_t	Jun_t	Sep_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Quarter 2	Aug_t	Sep_t	Dec_t	Sep_t	Sep_{t+1}	Sep_{t+2}
Quarter 3	Nov_t	Dec_t	Mar_{t+1}	Sep_{t+1}	-	Sep_{t+2}
Quarter 4	Feb_{t+1}	Mar_{t+1}	Jun_{t+1}	Sep_{t+1}	-	Sep_{t+2}

4. THE VOLUME INDEX IN GENERAL

The computation of the final monthly volume index of month m in year t is based on a wieghted Laspeyre's volume index, i.e.

$$I^m_{0t} = \sum_{i=1}^{n} w^m_{i0} \frac{Q^m_{it}}{Q^m_{i0}} = \sum_{i=1}^{n} \frac{P^m_{i0}Q^m_{i0}}{\sum_{i=1}^{n} P^m_{i0}Q^m_{i0}} \frac{Q^m_{it}}{Q^m_{i0}} = \frac{\sum_{i=1}^{n} P^m_{i0}Q^m_{it}}{\sum_{i=1}^{n} P^m_{i0}Q^m_{i0}} , \qquad (1)$$

where P^m_{it} and Q^m_{it} denote the price and the quantity, respectively, of commodity i in month m year t. Since the index concerns a compound of goods, a weight (w^m_{i}), measuring the particular commodity's share of the market value is used. This kind of wieghts is common in computations of price indexes (e.g. CPI), where the wieghts reflect the budget share of the commodity. However, the commodity's relative share of the market value does not necessarily reflect the relative share of resources used up in the production of it (Neter, Wasserman & Whitmore (1988), p 917). Other feasible wieghts could be hours worked or total costs, but such data are not available.

In the computation of the annual IPI a combination of weights is used. A Laspeyre's volume index is computed according to (1) for every plant, but in the aggregation of plants to branches the relative shares of the value added are used.

Another problem, which makes the use of (1) more difficult, is that the assortment of goods in the different markets changes as new goods replace old ones. The problem is the lack of a price for the base year. This problem is solved by constructing a chain index. An index for year t relative to t−1, $I^m_{(t-1),t}$, is computed according to (1). This index is then linked to the successive indexes of earlier years to achieve an index with the base year 0, i.e.

$$I^m_{0t} = I^m_{01} \cdot I^m_{12} \cdot \ldots \cdot I^m_{(t-2),(t-1)} \cdot I^m_{(t-1),t} \ . \qquad (2)$$

5. THE ANNUAL INDEX[2]

The preliminary versions of the monthly and quarterly indexes are based on the latest known final annual index (I_{0t}) for the branch in question. Therefore, we start the documentation of the methods of computation with a description of the computation of this index. The annual index of production (with base year 0) is created from the annual survey of the manufacturing industry in the following way:

[2]This chapter is a formalisation of the verbal description of the computations of the yearly indexes in SM:I, July 1991, p 2.

a) For every plant (g) a Laspeyre's volume index of production is computed for year t as compared to t−1, by dividing an index of the market value (PQ) in year t, (with base year t−1), with a Paasche's price index for the same period, i.e.

$$I^{g}_{t-1,t} = \frac{\sum_{i=1}^{n} (PQ)^{g}_{it}}{\sum_{i=1}^{n} (PQ)^{g}_{i(t-1)}} \frac{1}{\left(\frac{\sum_{i=1}^{n} (PQ)^{g}_{it}}{\sum_{i=1}^{n} P_{i(t-1)} Q^{g}_{it}} \right)} . \qquad (3)$$

$$= \frac{\sum_{i=1}^{n} P_{i(t-1)} Q^{g}_{it}}{\sum_{i=1}^{n} P_{i(t-1)} Q^{g}_{i(t-1)}}$$

The prices used in the computations are derived from the annual statistics. In these statistics there are data on the market value of annual production of every product in every plant $[(PQ)^{g}_{it}]$, as well as information on the quantity actually produced during that year (Q^{g}_{it}). From these data P_{it} may be computed as an average price for year t according to

$$P_{it} = \frac{\sum_{g} (PQ)^{g}_{it}}{\sum_{g} Q^{g}_{it}} . \qquad (4)$$

b) These volume indexes for each plant are weighed together to the lowest sub–branch level B_h (the 6 digit level in SNI). The weight of each plant corresponds to its relative share of the value added (X) of the sub–branch in year t−1.

$$I^{B_h}_{t-1,t} = \sum_g \frac{X^g_{t-1}}{\sum_g X^g_{t-1}} I^g_{t-1,t} \,, \tag{5}$$

where each plant g belongs to sub-branch B_h.

c) Using (5) in chain index formula (6) we get an annual index for t with base 0 for sub-branch B_h.

$$I^{B_h}_{0t} = I^{B_h}_{01} \cdot I^{B_h}_{12} \cdot \ \ldots \ \cdot I^{B_h}_{(t-2),(t-1)} \cdot I^{B_h}_{(t-1),t} \ . \tag{6}$$

d) The sub-branch indexes are aggregated to successively higher levels using the sub-branches' shares of the total value added of the branch (B) in the base year as weights, so that

$$I^{B}_{0t} = \sum_{B_h} \frac{X_0^{B_h}}{\sum_{B_h} X_0^{B_h}} I^{B_h}_{0t} \,, \tag{7}$$

where each sub-branch B_h belongs to branch B.

e) An aggregate index of the production volume (I_{0t}) for the manufacturing industry as a whole (SNI 3) is computed in the same way by weighing together the different branch indexes. The relative shares of the total value added in the base year serve as weights, i.e.

$$I_{0t}^{SNI3} = \sum_{B} \frac{X_o^B}{\sum_{B} X_0^B} I_{0t}^{B} \ . \qquad (8)$$

Further aggregation to an index for the industry as a whole (SNI 2+3) uses the same method. In order to facilitate the notations, we drop the branch notations (B och B_h), except when they are vital to understanding the text.

6. PRELIMINARY MONTHLY INDEXES[3]

The methods for computing the preliminary monthly indexes vary somewhat between branches, depending on what kind of indicator variable is used.

6.1 Branches with production data as indicators

In branches where data on production are directly available, the preliminary monthly index ($I^m_{A,0t}$) is computed from

$$I_{A,0t}^{m} = w_A^m I_{0f} = \frac{\sum_{i=1}^{n} P_{if} Q_{it}^{m}}{\left(\frac{1}{12} \sum_{m=1}^{12} \sum_{i=1}^{n} P_{if} Q_{if}^{m} \right)} I_{0f} \ , \qquad (9)$$

[3]The formalisation of the computational methods for monthly and quarterly indexes takes its starting point in the methodological description in Björkner (1988).

where P_{if} is the price of the last year for which final data are available, computed according to (6), and Q^{m}_{if} is the quantity of month m in year f, which is the last year for which final data are available.

This preliminary index is constructed as a chain index, where (the preliminary) monthly weight (w^{m}_{A}) provides the distribution of the annual index (I_{0f}) over the months. But this weight also serves as an index link from year f to year t. Note that w^{m}_{A} also may be interpreted as a Laspeyre's volume index for month m year t relative to the monthly average of year f. This index is then linked to the last final annual index I_{0f} to yield the development of production up to year t.

6.2 Branches with employment data as indicators

In branches where data on employment serve as indicators, the preliminary monthly index is computed as

$$I^{m}_{A,0t} = a^{m*}_{t} \rho^{*}_{t,q} \, , \qquad (10)$$

where a^{m*}_{t} denotes preliminary data on the number of hours worked in the branch during month m. The coefficient $\rho^{*}_{t,q}$ is an (intelligent) estimate of the development of productivity in that branch during the quarter. In connection with the quarterly revision (see Section 7) of the monthly index, $\rho_{t,q}$ is computed as the ratio between the preliminary quarterly index and the average number of hours worked during the quarter concerned.

6.3 Branches with delivery data as indicators

In the branches where delivery data are used as indicators the preliminary monthly index is cumputed through

$$I^{m}_{A,0t} = L^{m**}_{t}\,\delta^{*}_{t,q}\,. \tag{11}$$

Thus, the preliminary index is computed as the product of preliminary data on deliveries in fixed prices (L^{**}) and a factor ($\delta^{*}_{t,q}$), which corrects for the development of inventory investment, relative to deliveries, to yield a preliminary index of production.[4] In connection with the Q-revision of the monthly index, this correction factor is also revised and computed as the ratio between the quarterly index and the monthly average of deliveries during the quarter.

The preliminary data on deliveries in (11) are computed from

$$L^{m**}_{t} = \frac{l^{m**}_{h,t}\,u_{h,t}}{PPI^{m}_{h,t}} + \frac{l^{m**}_{ex,t}\,u_{ex,t}}{PPI^{m}_{ex,t}}\,, \tag{12}$$

where $l^{m**}_{h,t}$ and $l^{m**}_{ex,t}$ are preliminary data on deliveries in current prices in month m year t for the domestic market (h) and the export market (ex). These data on deliveries are deflated by the producer price index (PPI) for the same month, in the domestic and

[4]Given the basic relation between production (Q), deliveries (L), and inventory investment (ΔS), $Q = L + \Delta S$, δ may be expressed in the following way to emphasise the inventory correction aspect:

$$\delta = \frac{Q}{L} = \frac{L + \Delta S}{L} = 1 + \frac{\Delta S}{L}$$

export markets respectively, and are finally summed to yield the aggregate preliminary deliveries in fixed prices for the branch (L^{m**}_{t}).

When revised delivery data become available, a revision of the fixed price computation is made according to (12). Thus, the second version of delivery data (l^{m*}_{t}) makes it possible to compute a revised version of deliveries in fixed prices (L^{m*}_{t}). Finally, the final version of the deliveries in fixed prices (L^{m}_{t}) is computed from the final version of delivery data (l^{m}_{t}). The different revised versions of deliveries in fixed prices are then used in the computations of the following versions of the production index.

In Section 1 we mentioned that the data on deliveries are sample estimates. A sample of firms is used for a year and is then replaced. In order to make the computations of IPI consistent and independent of the changes in the sample, sample factors for the domestic and export markets ($u_{h,t}$ and $u_{ex,t}$ respectively in (12)) are computed from a double set of data. Through the survey design, which was described in Section 1, access to a double set of data for November and December is obtained, since the January enquiry to the new sample of firms also yields data for November and December the previous year. These data can be compared to the November and December data from the old sample of firms. From this double set of data the sample factor for the export market is computed as

$$u_{ex,t} = \frac{\sum_{m} l^{m,old}_{ex,t-1}}{\sum_{m} l^{m,new}_{ex,t-1}} \quad ; \; m = Nov \wedge Dec \; year \; t-1 \; . \qquad (13)$$

The computation of the sample factor for the domestic market is analogous to (13).

7. PRELIMINARY QUARTERLY INDEXES

The computation of the quarterly index, and later the Q– and Y–revisions of the monthly indexes, for a branch are based on the rules of consistency that have to prevail between the final monthly ($I^m_{F,0t}$), quarterly ($I^q_{F,0t}$), and annual (I_{0t}) indexes (Öhlén (1991), p 7), namely:

1. Consistency between the monthly and annual indexes:

$$\sum_{m=1}^{12} I^m_{F,0t} = 12\, I_{0t} \ . \tag{14}$$

2. Consistency between the monthly and quarterly indexes:

$$\sum_{m=1}^{3} I^m_{F,0t} = 3\, I^q_{F,0t} \ . \tag{15}$$

3. Consistency between the quarterly and annual indexes:

$$\sum_{q=1}^{4} I^q_{F,0t} = 4\, I_{0t} \ . \tag{16}$$

7.1 Quarterly indexes based on production data

In branches where data on actual production are available as inputs, the data that are used for computing the preliminary monthly index are simply summed over the quarter to obtain the quarterly index. This procedure differs from that in other branches, where additional complementary data are used.

Using the rule of consistency between monthly and quarterly indexes (15), the preliminary quarterly index is computed from

$$I^{q}_{A,0t} = \frac{\sum_{m=1}^{3}\sum_{i=1}^{n} P_{if}Q^{m}_{it}}{\left(\frac{1}{4}\sum_{m=1}^{12}\sum_{i=1}^{n} P_{if}Q^{m}_{if}\right)} I_{0f} \; . \qquad (17)$$

7.2 Quarterly indexes based on employment or delivery data

In branches where data on production are not available as inputs, the first preliminary version of the quarterly index can be computed using (15) as soon as preliminary monthly data for the quarter become available. Summing of the monthly data across the months (m_1, m_2 and m_3) of the quarter yields the quarterly index.

For branches where employment data are used as inputs the preliminary quarterly index is computed as

$$I^{q}_{A,0t} = \frac{1}{3}\left(a_t^{m_1} + a_t^{m_2} + a_t^{m_3 *}\right)\rho^{*}_{t,q} . \qquad (18)$$

Employment data for the last month of the quarter are preliminary in this computation. For the first two months final data are available. For branches where delivery data are used as indicators, the preliminary quarterly index is computed from

$$I^{q}_{A,0t} = \frac{1}{3}\left(L_t^{m_1} + L_t^{m_2 *} + L_t^{m_3 **}\right)\delta^{*}_{t,q} . \qquad (19)$$

At the time of this computation final delivery data are available for the first month of the quarter only. In these branches complementary data on deliveries and inventory investment are used to compute the later index versions. In order to distribute the annual index I_{0t} over the quarters, so called production estimates based on the delivery and inventory statistics from Statistics Sweden are produced. The production estimate ($e_{t,q}$) for quarter q year t for the branch in question is computed as the sum of deliveries in fixed prices and the change in inventories during quarter q year t, i.e.

$$e_{t,q} = \sum_{m=1}^{3} L_t^{m} + \Delta S_{t,q} . \qquad (20)$$

Using the rule of consistency (16) the final quarterly index is computed as

$$I^{q}_{F,0t} = \frac{e_{t,q}}{\sum_{q=1}^{4} e_{t,q}} 4\, I_{0t} . \qquad (21)$$

Adding both sides of (21) for all quarters in year t shows that the rule of consistency holds. Thus, (21) is the formula for computing the final quarterly index. Since the annual industry statistics have a time lag of 15–18 months, the final quarterly index can not be produced until year t+2 (Table 2). Furthermore, there is a certain time lag before data on deliveries and inventory investment become available.

Two and a half months after the end of a quarter preliminary inventory data become available. This makes it possible to find a preliminary estimate of production through

$$e^{*}_{t,q} = \left(L_t^{m_1} + L_t^{m_2} + L_t^{m_3 *}\right) + \Delta S^{*}_{t,q} . \qquad (22)$$

Data on deliveries for the first two months are final, but data on inventory investment and on deliveries for the last month are preliminary.

The latest annual index known is from year f, I_{0f}. In order to find a consistent preliminary quarterly index (version 2) the production estimates from the last year for which final data are available are used. Thus, the preliminary quarterly index is computed as

$$I^{q}_{B,0t} = \frac{e^{*}_{t,q}}{\sum_{q=1}^{4} e_{f,q}} 4\, I_{0f} , \qquad (23)$$

which is used in the first Q–revision of the monthly index.

The third version of the preliminary quarterly index can be computed as soon as final data on deliveries and inventories become available. A final production estimate is then found using (20), and the third quarterly index version is computed according to

$$I^q_{C,0t} = \frac{e_{t,q}}{\sum_{q=1}^{4} e_{f,q}} 4\, I_{0f} \; . \tag{24}$$

This version of the quarterly index is used in the second Q-revision of the monthly index. It is worth noting that the second and third versions of the quarterly index are based on data on deliveries and inventories, even in branches where employment data are used as indicators.

8. Q-REVISIONS OF THE PRELIMINARY MONTHLY INDEXES

In the branches where data on production are used in the computations of the indexes, no Q-revisions of the preliminary monthly indexes are made. In other branches the Q-revisions are based on the rule of consistency (15) between monthly and quarterly indexes.

8.1 Q-revisions of the monthly indexes for branches where employment data are used as indicators

The Q-revision of the preliminary monthly index is made as soon as the productivity factor $\rho^*_{t,q}$ becomes available. This in turn happens when the preliminary quarterly index (version 2) is published. The first Q-revision of the monthly index can be described as

$$I^m_{B,0t} = a^m_t \rho^{rev1}_{t,q} \; . \tag{25}$$

The derivation of $\rho^{rev1}_{t,q}$ uses the rule of consistency (15). The factor $\rho^{rev}_{t,q}$ is given by

$$\sum_{m=1}^{3} I^{m}_{B,0t} = \rho^{rev1}_{t,q} \sum_{m=1}^{3} a^{m}_{t} = 3\, I^{q}_{B,0t}$$

$$\rho^{rev1}_{t,q} = \frac{3\, I^{q}_{B,0t}}{\sum_{m=1}^{3} a^{m}_{t}} \qquad (26)$$

Thus, the first Q–revision of the monthly index is given by

$$I^{m}_{B,0t} = a^{m}_{t} \rho^{rev1}_{t,q} = a^{m}_{t} \frac{3\, I^{q}_{B,0t}}{\sum_{m=1}^{3} a^{m}_{t}} , \qquad (27)$$

where $I^{q}_{B,0t}$ is given in (23).

When the third version of the quarterly index has been computed as in (24), a second Q–revision of the monthly index is made. (This means that the productivity factor ρ is once again revised.) The computation of this version of the monthly index is analogous to (27), i.e.

$$I^{m}_{C,0t} = a^{m}_{t} \rho^{rev2}_{t,q} = a^{m}_{t} \frac{3\, I^{q}_{C,0t}}{\sum_{m=1}^{3} a^{m}_{t}} . \qquad (28)$$

8.2 Q-revisions of the monthly index for branches where delivery data are used as indicators

In branches where the index is based on delivery data, the Q-revision is made in the same way as in branches where employment data are used. Using (15) the inventory correction factor for these branches ($\delta^{*}_{t,q}$) can be computed as soon as the preliminary quarterly index (version 2) is known. The first Q-revision of the monthly index in these branches is made according to

$$I^{m}_{B,0t} = L^{m}_{t}\,\delta^{rev1}_{t,q} = L^{m}_{t}\,\frac{3\,I^{q}_{B,0t}}{\sum_{m=1}^{3} L^{m}_{t}} \;. \qquad (29)$$

The second Q-revision takes place as soon as the third version of the quarterly index becomes available, i.e.

$$I^{m}_{C,0t} = L^{m}_{t}\,\delta^{rev2}_{t,q} = L^{m}_{t}\,\frac{3\,I^{q}_{C,0t}}{\sum_{m=1}^{3} L^{m}_{t}} \;. \qquad (30)$$

9. Y-REVISIONS. FINAL MONTHLY AND QUARTERLY INDEXES

The Q-revised monthly indexes and the preliminary quarterly indexes are based on the last final annual index known (I_{0f}) for the branch. Access to new final annual data therefore makes Y-revisions of the different monthly and quarterly indexes possible.

In order to describe the pattern of Y-revisions of a monthly (I^m_{0t}) or quarterly index (I^q_{0t}) in a more formal way, we may express the last year for which final data are available as f=t−j, where t is the year that the index concerns and j gives the time lag of the industrial statistics. For months and quarters in the first half of the year j=3 and for months and quarters in the second half of the year j=2. The annual index $I_{0,f}$, may now be written $I_{0,t-j}$. When final data for the next year become available f=t−j+1. The first Y-revision lets $I_{0,t-j+1}$ replace $I_{0,t-j}$ in the computations of the monthly and quarterly indexes. Y-revisions are made until f=t. The number of Y-revisions depends on what half of the year the index concerns.

The final monthly and quarterly indexes and the final annual index must be related to each other according to the rules of consistency (14) − (16).

9.1 Y-revisions in branches where production data are used as indicators

In branches where preliminary indexes are based on data on production, the monthly indexes are revised by updating the information on prices and quantities in (9) using the last annual data known. Thus, the first Y-revision is made according to

$$I^m_{D,0t} = w^m_D I_{0,(t-j+1)} = \frac{\sum_{i=1}^{n} P_{i,(t-j+1)} Q^m_{it}}{\left(\frac{1}{12}\sum_{m=1}^{12}\sum_{i=1}^{n} P_{i,(t-j+1)} Q^m_{i,(t-j+1)}\right)} I_{0,(t-j+1)} \quad . \tag{31}$$

Note that the month-weights (w^m) are updated and still work as index links from the last year with final data to the current year t. The final monthly index is computed from

$$I_{F,0t}^{m} = w_F^m I_{0t} = \frac{\sum_{i=1} p_{it}q_{it}}{\left(\frac{1}{12}\sum_{m=1}^{12}\sum_{i=1}^{n} p_{it}Q_{it}^{m}\right)} I_{0t} \; . \tag{32}$$

In the final computation the month–weights only perform the function of distributing the final annual index over the months of the year. The Y–revisions of the quarterly indexes are made using (31) and (32) and summing across the months of the respective quarter.

9.2 Y–revisions for branches where employment data are used as indicators

When the annual index of the year t–j+1 becomes available, the first Y–revision of the quarterly index can be made according to

$$I_{D,0t}^{q} = \frac{e_{t,q}}{\sum_{q=1}^{4} e_{(t-j+1),q}} 4\, I_{0,(t-j+1)} \; , \tag{33}$$

$e_{t,q}$ being the production estimate derived in equation (20). This revised quarterly index is then used in the Y–revision of the monthly index. At the same time the productivity factor of the branch is once again revised. The revision is made according to

$$I_{D,0t}^{m} = a_t^m \rho_{t,q}^{rev} = a_t^m \frac{3\, I_{D,0t}^{q}}{\sum_{m=1}^{3} a_t^m} = \frac{a_t^m}{\sum_{m=1}^{3} a_t^m} \frac{e_{t,q}}{\sum_{q=1}^{4} e_{(t-j+1),q}} 12\, I_{0,(t-j+1)} \; . \tag{34}$$

The Y–revisions proceed according to this pattern, until the annual industrial statistics become available for that year, i.e. until f=t. The final quarterly index is given by (21), while the final monthly index is computed from

$$I_{F,0t}^{m} = a_t^m \rho_{t,q}^{fin} = a_t^m \frac{3\, I_{F,0t}^{q}}{\sum_{m=1}^{3} a_t^m} = \frac{a_t^m}{\sum_{m=1}^{3} a_t^m} \frac{e_{t,q}}{\sum_{q=1}^{4} e_{t,q}} 12\, I_{0t} , \qquad (35)$$

which means that the rule of consistency between monthly and annual indexes holds.[5]

9.3 Y-revisions for branches where delivery data are used as indicators

For branches where the indexes are based on input data on deliveries, the same methods are used in the Y-revisions, and in the final computations of the quarterly index, as for the branches where the indicators are employment data. As for the monthly index the methods are completely analogous and the first Y-revision of the monthly index is made according to

$$I_{D,0t}^{m} = L_t^m \delta_{t,q}^{rev} = L_t^m \frac{3\, I_{D,0t}^{q}}{\sum_{m=1}^{3} L_t^m} . \qquad (36)$$

The final version of the monthly index in these branches is computed from

[5]To see that (35) is in accordance with the rule of consistency, note that the sum of all m belonging to year t may be divided into a double sum as in

$$\sum_{m=1}^{12} = \sum_{m=1}^{3} \sum_{q=1}^{4} .$$

$$I^{m}_{F,0t} = L^{m}_{t}\delta^{fin}_{t,q} = L^{m}_{t}\frac{3\,I^{q}_{F,0t}}{\sum_{m=1}^{3} L^{m}_{t}} = \frac{L^{m}_{t}}{\sum_{m=1}^{3} L^{m}_{t}}\,\frac{e_{t,q}}{\sum_{q=1}^{4} e_{t,q}}\,12\,I_{0t}\;. \qquad (37)$$

The ways in which the Y–revisions are made imply that the final monthly and quarterly indexes are based on a combination of sample based indicators and total survey data (q^{m}_{it}, a^{m}_{t} and L^{m}_{t} are all estimates based on sample data).

9.4 A comment on the nature of revisions

The revisions that are performed do not influence the relationship between the monthly indexes within a given quarter or the relationship between the quarterly indexes within a given year. In branches where employment data are used as indicators, this means that the relations between the indexes of any two quarters q_1 and q_0 in year t are constant from the third preliminary version on, and may be written

$$\frac{I^{q_1}_{C,0t}}{I^{q_0}_{C,0t}} = \frac{I^{q_1}_{D,0t}}{I^{q_0}_{D,0t}} = \frac{I^{q_1}_{E,0t}}{I^{q_0}_{E,0t}} = \frac{I^{q_1}_{F,0t}}{I^{q_0}_{F,0t}} = \frac{e_{t,q_1}}{e_{t,q_0}}\;. \qquad (38)$$

In these branches the relation between any two months m_1 and m_0 in year t may be written

$$\frac{I^{m_1}_{C,0t}}{I^{m_0}_{C,0t}} = \frac{I^{m_1}_{D,0t}}{I^{m_0}_{D,0t}} = \frac{I^{m_1}_{E,0t}}{I^{m_0}_{E,0t}} = \frac{I^{m_1}_{F,0t}}{I^{m_0}_{F,0t}} = \frac{a^{m_1}_t}{a^{m_0}_t} \frac{\sum_{m=1}^{3} a^m_t}{\sum_{m=1}^{3} a^m_t} \frac{e_{t,q_1}}{e_{t,q_0}} , \qquad (39)$$

where m_0 and m_1 belong to the quarters q_0 and q_1 respectively. If the months m_0 and m_1 belong to the same quarter, the monthly relations in (39) are reduced to a^{m1}_t/a^{m0}_t. In branches where delivery data are used as indicators, the relations between the indexes are analogous to those in branches where employment data are used.

Note also that (38) and (39) presuppose that no revisions of employment data over and above the "normal" revisions are made and that these constant index realations apply from version C of the index on. In this version final sample data on hours worked, deliveries, and inventory investment are used.

In the branches where production data are used as indicators, the Y-revisions may imply changes in the relations between indexes as well as in their levels, since new data on both prices and quantities are used.

10. A DESCRIPTIVE STUDY OF MONTHLY INDEXES

Here the relationships between the revised versions of IPI and the final index are studied, in particular two hypotheses:

1: The process of measurement is under control.

If the process of measurement is under control, there should be no systematic errors in the estimations, i.e. the hypothesis that the mean deviation equals zero should not be rejected. There should be no positive or negative trend in the measurment errors, and the variability should not increase over time.

2: The accuracy improves after revision.

A reasonable assumption is, that early publication has a cost in the form of uncertainty, i.e. that the accuracy, relative to the final figure, is lower the earlier a revision is made. This implies that both the mean deviation and the variability should decrease with every revision. We use the Standard Deviation and the Mean Absolute Deviation (MAD) to measure the variability.

We have investigated the manufacturing industry as a whole (SNI 3) and the machinery manufacturing (SNI 38). Our results cannot be generalised for the measurement process as a whole, since we do not formally test the hypotheses above, but only describe a part of the process. Nevertheless, the results of our study raise doubts about whether the measurment process is under control in a statistical sense.

10.1 Outline of the investigation

The data series have been taken from Statistical Reports, Series I. The following series have been used (See Section 3.2): The preliminary index (A), Q–revision 1 (B), Y–revision 1 (D), and the final index (F). 1970 is the base year in all series. The more recent parts of the series, that have been published with 1980 as a base year, have been recomputed with the link that was used at the first publication after the change. Some small errors may have arisen because all indexes have been rounded off to integer values. In the aggregate manufacturing industry the 192 observations between 1973:1

and 1988:12 have been used, and in the machinery manufacturing the 204 observations between 1972:1 and 1988:12.

10.2 Results for the machinery manufacturing industry

The preliminary index as well as the revised versions tend to be overstatements compared to the final index, as can be seen in Table 1.

Table 3. A comparison of the accuracy of the different versions of IPI for machinery manufacturing.

	MEAN	STDEV	MAD	MEDIAN	Q1	Q3
Prel-Final	2.7	2.5	4.6	2.5	-1.0	6.0
Q-rev-Final	2.5	3.9	3.5	2.5	0.0	6.0
Y-rev-Final	0.9	3.9	2.3	1.0	-1.0	3.0

The preliminary index exceeds the final index by 2.7 index units on average. 54% of all indexes overstate, 21% understate, and 25% are equal to the final index. The overstatements are not concentrated to any particular time period. However, some special events like strikes, oil crises, and devaluations coincide with large deviations between preliminary and final numbers. The three largest deviations in the sample occured in the same quarter, namely 1976:1. These overstatements were alreadey corrected in the first Q–revision.

The first Q–revision also yields some systematic overstatement, the mean deviation being 2.5 index units. There seems to be no systematic difference between the quarters in this case, however. The mean deviation after the first Y–revision is 0.9 index units, and no systematic difference between the quarters can be found. MAD decreases as new revisions are performed. MAD is 4.6 index units in the preliminary version, 3.5 in the first Q–revision, and 2.3 in the first Y–revision. The biases in the different index versions relative to the final one are shown graphically in Figure 1.

Figure 1. The distributions of the accuracy of the different versions of IPI.

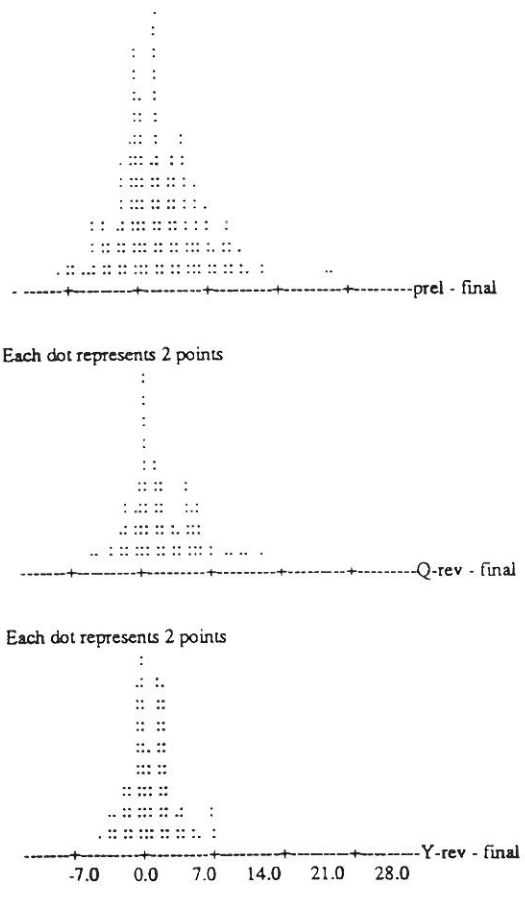

Table 2 compares the deviations between the preliminary and the final series over the quarters. The overstetement is most prominent in the first quarter, 5.6 index units on average. The preliminary index overstates in 77% of all cases and understates in only 4%. The mean overstatement in the quarters is between one and two index units.

Table 4. A comparison of the deviations between the preliminary and final series and the distribution of quarterly over- and underestimates for the machinery manufacturing.

	MEAN	STDEV	MAD	OVER ESTIM	UNDER ESTIM	"CORRECT"
Quarter 1	5.6	5.8	5.8	77%	4%	29%
Quarter 2	1.5	4.8	4.0	50%	25%	25%
Quarter 3	1.6	5.6	4.0	40%	25%	35%
Quarter 4	2.0	5.2	4.5	56%	22%	22%

The variability within the quarters is virtually equal in all four quarters, the standard deviation being between five and six. MAD is 5.8 in the first quarters and about four in the other quarters of the year.

10.2 Results for the aggregate manufacturing industry

On the aggregate level the large systematic differences that we found on the branch level have disappeared, as Table 5 shows.

Table 5. A comparison of the accuracy between the preliminary and the final IPI–series in the aggregate manufacturing industry.

	MEAN	STDEV	MAD	MEDIAN	Q1	Q3
Prel.–Final	0.8	3.0	2.4	1.0	–1.0	3.0

In combination with the strong tendency for the preliminary index to overstate the final index in the machinery manufacturing, this implies that the preliminary indexes for some of the other branches are systematically understated.

11. CONCLUDING REMARKS

The most striking characteristic of the production process behind the index of industrial production is the complexity of it. The preliminary index is computed from several different sets of data, collected from different sources and with different methods. The preliminary index is then revised several times using additional data collected from other sources. Any change in an annual or quarterly index causes a chain reaction, bacause of the need for consistency between monthly, quarterly, and annual indexes. In most branches the input data are sample estimates of deliveries, employment, and inventory investment, which obviously means that there is a sample variation in the index. Furthermore, the plants in the samples are substituted regularly, which decreases the comparability between index numbers over time. These sources of uncertainty call for the use of confidence intervals rather than point estimates, whenever the index is used in forecasts or analyses. The proper confidence intervals are of course difficult to compute, but obviously some effort should be made to determine what variations in the index are reasonably significant.

Hopefully, this paper can help to make users of the index of industrial production more aware of the uncertainties connected with it. It is obvious, that this uncertainty carries over to any forecast or analysis that is based on the index. Also, this paper may form a basis for any future attempt to improve the methods behind the index, or to compute a confidence interval for it.

12. ACKNOWLEDGEMENTS

In particular we want to thank Mickael Löthgren, who in connection with his graduation work has participated in the work on this report. A great deal of the work on collecting material as well as important parts of the mathematical formalisations have been made by him. Lars–Erik Öller has given valuable comments on an earlier draft. Also, we want to offer our thanks to Sven Näsman at Statistics Sweden for his assistance. Support from the Economic Council of Sweden is gratefully acknowledged. Remaining shortcomings of the report are of course the responsibility of the authors alone.

13. SAMMANFATTNING

Industriproduktionsindex, IPI, produceras av SCB på såväl månads- som kvartals- och årsbasis och publiceras i SM, Serie I. Månads- och kvartalsindex publiceras i vardera minst sex versioner, från den första preliminära via ett antal revisioner till den slutliga definitiva versionen. Indexen baseras på produktionsdata från olika branschorganisationer, sysselsättningsdata eller leveransdata från SCB:s korttidsstatistik, beroende på vilken av industrins delbranscher det gäller.

Det första preliminära månads- eller kvartalsindexet publiceras med en eftersläpning på två månader, medan det definitiva indexet dröjer ca två år. Årsindexet baseras på en totalundersökning, medan månads- och kvartalsindexen bygger på urvalsundersökningar. Årsindexet utnyttjas vid beräkningen av kvartalsindexet, varför detta revideras då nya årsindex blir tillgängliga. Därutöver revideras kvartalsindexet då säkrare inputdata kan användas. Kvartalsindexet utnyttjas i sin tur vid beräkningen av månadsindexet, varför detta revideras då nya versioner av kvartalsindexet blir tillgängliga. Relationerna mellan månadsindexen inom respektive kvartal revideras emellertid inte. Denna uppsats ger för första gången en formell beskrivning av beräkningen av de olika versionerna av månads- och kvartalsindexen.

Uppsatsen avslutas med en kontroll av träffsäkerheten i det första preliminära, det första kvartalsreviderade och det första årsreviderade månadsindexet jämfört med det definitiva, dels för den aggregerade tillverkningsindustrin, dels för verkstadsindustrin. I synnerhet den preliminära, men även de reviderade versionerna visar en systematisk överskattning i förhållande till det definitiva indexet i verkstadsindustrin, men inte på aggregerad nivå, vilket indikerar att det finns systematiska underskattningar i någon eller några andra delbranscher i industrin.

REFERENCES

Björkner, E, (1988), "The Production Volume Index" (In Swedish: "Produktionsvolymindex (PVI)"), Mimeograph, Statistics Sweden, F/STM.

Neter, J, Wasserman, W & Whitmore, G A, (1988), "Applied Statistics", 3rd ed, Allyn and Bacon Inc, Newton, Massachusetts.

Skalin, J, (1990), "Notes on Forecasts and Revisions of Preliminary Swedish Production Statistics", Mimeograph, Stockholm School of Economics.

Statistical Reports, Series I (SM:I), "Index of Industrial Production, July 1991" (In Swedish: "Industriproduktionsindex för Juli 1991"), Statistics Sweden, Department of Business Cycle Statistics.

Tengblad, Å, (1992), "The Reliability of the Quarterly National Accounts" (In Swedish: "Tillförlitligheten i de svenska kvartalsvisa nationalräkenskaperna"). Arbetspapper Nr. 2, April 1992. National Institute of Economic Research.

Öhlén, S, (1991), "Revision of the Short-Term Statistics: Optimal Level Adjustment of the Monthly Index of Industrial Production" (In Swedish: "Översyn av den Kortperiodiska Statistiken: Optimal nivåanpassning av månatlig industriproduktionsindex"), Preliminary report, Statistics Sweden, F/STM.

Prevoius titles in this serie:

No. 1 Current Account and Business Cycles: Stylized Facts for Sweden by Anders Warne and Anders Vredin. December 1989.

No. 2 Change in Technical Structure of the Swedish Economy by Göran Östblom. December 1989.

No. 3 Mamtax. A Dynamic CGE Model for Tax Reform Simulations by Paul Söderlind. December 1989.

No. 4 The Supply Side of the Econometric Model of the NIER by Alfred Kanis and Aleksander Markowski. November 1990.

No. 5 The Financial Sector in the SNEPQ Model by Lennart Berg. February 1991.

No. 6 Consumer Attitudes, Buying Intentions and Consumption Expenditures. An Analysis of the Swedish Household Survey Data by Anders Ågren & Bo Jonsson. April 1991.

No. 7 A Quarterly Consumption Funtion for Sweden 1979-1989 by Lennart Berg & Reinhold Bergström. October 1991.

No. 8 Good Business Cycle Forecasts - A Must for Stabilization Policies by Lars-Erik Öller. February 1992.

No. 9 Forecasting Car Expenditures Using Household Survey Data by Bo Jonsson and Anders Ågren. February 1992.

No. 10 Forecasting the Business Cycle Not Using Minimum Autocorrelation Factors by Karl-Gustaf Löfgren, Bo Ranneby & Sara Sjöstedt. February 1992.

No. 11 Current Quarter Forecasts of Swedish GNP Using Monthly Variables by Stefan Gerlach. February 1992.

No. 12 The Relationship Between Manufacturing Production and Different Business Survey Series in Sweden by Reinhold Bergström. February 1992.

No. 13 Forecasting the Swedish Unemployment Rate: V A R vs Transfer Function Modelling by Per-Olov Edlund and Sune Karlsson. March 1992.

No. 14 Business Survey Data in Forecasting the Output of Swedish and Finnish Metal and Engineering Industries: A Kalman Filter Approach by Markku Rahiala and Timo Teräsvirta. March 1992.

No. 15 The Relationship Between Manufacturing and Various BTS Series in Sweden Illuminated By Frequency and Complex Demodulate Methods by Anders Christoffersson, Roland Roberts and Ulla Eriksson. April 1992.

No. 16 Sample Based Proportions As Values On An Independent Variable In A Regression Model by Bo Jonsson. October 1992.

No. 17 Eliciting Turning Point Warnings From Business Surveys by Lars-Erik Öller. November 1992.

No. 18 Volatility, Trading Mechanisms and International Cross-Listing by Margaret M. Forster. December 1992.

No. 19 Prediction With a Linear Regression Model and Errors in a Regressor by Bo Jonsson. December 1992.

No. 20 Corporate Control, Portfolio Choice, and the Decline of Banking by Gary Gorton and Richard Rosen.