Introducing a new method to calculate index weights for the Producer Price Indices

Louise Morris and Jon Gough
Business Prices and Sales Division
Office for National Statistics
Room 1.162 Government Buildings
Cardiff Road
Newport, NP10 8XG
Tel: 01633 813374
E-mail: jon.gough@ons.gov.uk

Summary

- The Producer Price Index (PPI) is currently being rebased which means updating the weighting structure from the base year 1995=100 to 2000=100. As a part of this process a review of certain methodological aspects has been undertaken.

- One such change was to look at the approach to the calculation of index weights and adopt a more sophisticated method to estimate sales to the domestic market, used to weight detailed PPI series to produce indices at a more aggregate level.

- Introducing this new method of determining weights will only lead to small differences in index values at the aggregate level with larger differences confined to more detailed indices.

- The rebased results are to be published in October 2003. At that point there will be further changes due to the updating of the weights.

Introduction

Background to Producer Price Inquiry (PPI)

The PPI measures the change in prices of goods bought and sold by UK manufacturers. Overall there are four types of PPI series produced, which are:

- Gross Sector Output (GSO);
- Net Sector Output (NSO);
- Gross Sector Input (GSI);
- Net Sector Input (NSI).

With the difference between output and input being:

- Output prices – measure the change in price of goods sold by UK manufacturers;

- Input prices – measure the change in price of goods bought by manufacturers for use in the manufacturing process.

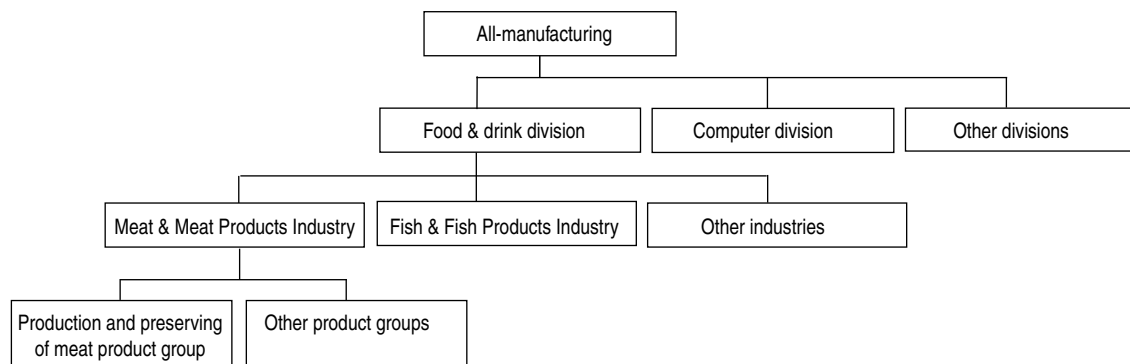And the difference between net and gross sector is:

- Net sector – the weights used to calculate these exclude transactions between companies classified to the same sector, e.g. the value of an electronic component manufacturer's sales to a car manufacturer would be excluded from the weights (thereby reflecting the value of sales to purchasers outside the manufacturing sector);

- Gross sector – all transactions are included when deriving the weights, including sales within the same sector.

The same basic price information is used to feed into each of these four types of PPI series. The difference between the various indices is the weights that are applied to combine the low level series to form these higher level indices and which low level series are combined to form the high level indices. The headline series produced in the PPI First Release are the NSO and NSI all-manufacturing series including duty.

Price Data

Around 9,000 price quotes are collected each month together with some prices from administrative sources like trade publications and other government departments. Output PPIs are calculated at a fairly detailed product group (six digit) level, with the products that fall into each PPI defined by the European 'Classification of Products by Activity' (CPA) which in turn is based on the 1992 Standard Industrial Classification. Indices produced for 1,277 detailed product groups are then grouped together using the 'family tree' structure of the CPA to produce 229 industry (four digit) level series. The industry level series are then grouped to give 23 division level (two-digit) indices, which in turn are grouped into the 'all-manufacturing' index. An example of this structure is provided in Figure 1.

Figure 1

```
                          ┌──────────────────┐
                          │ All-manufacturing │
                          └──────────────────┘
              ┌────────────────────┼────────────────────┐
    ┌──────────────────┐  ┌──────────────────┐  ┌──────────────────┐
    │ Food & drink      │  │ Computer division │  │ Other divisions   │
    │ division          │  │                   │  │                   │
    └──────────────────┘  └──────────────────┘  └──────────────────┘
      ┌────────────┼────────────────┐
┌──────────────────┐ ┌──────────────────┐ ┌──────────────────┐
│ Meat & Meat       │ │ Fish & Fish       │ │ Other industries  │
│ Products Industry │ │ Products Industry │ │                   │
└──────────────────┘ └──────────────────┘ └──────────────────┘
   ┌──────┴──────┐
┌──────────────────┐ ┌──────────────────┐
│ Production and    │ │ Other product     │
│ preserving of     │ │ groups            │
│ meat product group│ │                   │
└──────────────────┘ └──────────────────┘
```

Index weights

The high level all-manufacturing series are generally structured in the same way. It is easiest to explain the structure for the GSO. Initially the prices supplied by each contributor are compared to the average price of the same item in the base period to form a price relative. The price relatives are then weighted together to form the 6-digit product index. The weights are derived based on the value of ProdCom sales (total sales figure obtained from the Products of the European Community survey). The 6-digit product groups are then weighted together with other product groups of a similar nature to produce the industry indices. In turn these industries are weighted together to form their respective divisional indices. Finally weighting together all the divisional series then produces the GSO all-manufacturing index.

Indices from product level to divisional level are produced on a gross sector basis. At the all-manufacturing level output indices are produced on both a gross and net sector basis. To calculate the NSO series the same method is used to produce indices from product level to generally divisional level as is used for the GSO series. To combine the division level indices to produce the all-manufacturing NSO series, input-output data is used in place of ProdCom (and export) data to provide index weighting patterns. Unlike ProdCom data which provides only a total product sales value input-output data allows a split in sales to be made within and outside of the manufacturing sector, enabling sales to the manufacturing sector to be excluded from the NSO weights.

The Net Sector Input series is calculated from import and gross sector output indices which are calculated up to input/output group level using similar methods to those described above. These series are then weighted together using input/output domestic and import data, removing sales and imports to the manufacturing sector, in the same way as for NSO.

Rebasing

Rebasing is a five- yearly process for the PPI (and trade price series)

with the aim being to update the weights used in the calculation of all series produced. Over time relative volumes and prices of products sold will change and it is important to ensure the weighting structure of the index is updated at regular intervals to reflect recent information on the relative importance of products

.

Together with the updating of the weights to the most recent base year other methodological issues have been addressed during the current rebasing exercise. There are three key methodological issues:

1. The change to the calculation of the sales to the domestic market (total sales less exports) data which is used to weight the output indices for the PPI;
2. The weighting pattern and structure of the import series;
3. The structure of the input series.

This article will look mainly at the change in method to calculate the home sales data and the effect this has had on the actual results. Brief summaries are detailed towards the end of the report for the other changes implemented as part of the rebasing process.

Calculation of home sales data

Current method

Each 6-digit product level index is produced by weighting together price information supplied by a sample of contributors. For further details on the approach to index calculation see the *Economic Trends* article (Morris and Birch, 2001).

To move from the 6-digit to higher level series, indices are weighted together based on the relative values of sales of each 6-digit product group to the domestic (home) market. These values used as the basis for index weights are referred to here as 'home sales'.

Home sales are calculated using ProdCom total sales estimates (which include sales to both the domestic and export markets) adjusted using export data from Customs and Excise (C&E) to

remove exports from the total, leaving just the required estimates of sales to the home market. The weights of relevant product level indices within the industry level index are the home sales value for the 6-digit product group divided by the total home sales of all 6-digit product groups within the industry.

Whilst in theory it is a straightforward calculation to produce home sales values, problems are encountered as the data is obtained from two different sources. There are differences in the approach used to derive the respective estimates of totals produced, both in terms of methods and also in terms of classification and definition. (See the section 'Reason for changing the method' for further details). Cases were found where the export value for particular product groups were greater than the total sales. In these cases 'special rules' are adhered to so all negative home sales are removed. An adjustment to remove negative home sales estimates was also applied in previous rebasing exercises. For the current exercise an improvement has been made to the method used.

Reason for changing the method

The review of the approach taken to adjusting these negative estimates addressed concerns with the previous approach namely:

1. Matching two data sources.

The ProdCom estimates of total sales and Customs and Excise estimate of export sales are derived using different methods, coverage and definitions – there were only limited allowances for these differences in place.

ProdCom is a survey of businesses classified to the manufacturing sector. A stratified random sample design is used with stratification based on industry classification of the business and employment. All of the larger businesses and a sample of the smaller businesses within each industry are selected for the survey. Contributors are then asked to provide information on sales of products manufactured. Product classifications are based on the European ProdCom 'Product list'. In contrast Customs & Excise estimates of product exports are based on a cut-off sample of traders with significant total exports covering around 97 per cent of total trade. Product classifications are based on the 'combined nomenclature' (CN) classification system.

There could therefore be differences between the two data sources due to:

- Coverage – ProdCom includes only businesses classified to the manufacturing sector whilst customs covers all exports (including re-exports).
- Design – there are likely to be differences in responders to the survey.

- Classification – products may be classified differently by responders (although there is a direct link between ProdCom and CN classifications).

2. There may be some bias in the approach previously used as only negative values were adjusted and not all values were adjusted in a similar way. The magnitude of this bias was determined as the value of exports of unrecorded production (e.g. re-exports and production from industries not included in ProdCom).

New method

Initially, in developing the revised approach, consideration was given to whether additional information was available to supplement the sales estimates available from ProdCom and C&E surveys. The monthly Production Inquiry (MPI) was identified as an additional source of information. This survey provides monthly estimates of total and export sales at an industry level. Whilst there are differences between the MPI and PPI approach, MPI estimates are based on industry classifications whilst PPI estimates are produced on a product classification basis; MPI data provided a useful additional source of information.

A more detailed but robust method to produce home sales estimates has now been defined which addresses concerns with the previous approach. The underlying rationale for the improved method is that C&E data is used to provide relative export proportions by product group. These are applied to each ProdCom industry separately, to obtain a split between exports and home sales. These estimates are then scaled to ensure that the ProdCom industry export totals are consistent with the MPI data.

Average C&E export proportions are compared against the MPI export proportion. If the C&E export proportion is greater then exports are defined as too big and are reduced. If the C&E export proportion is less than the MPI export proportion then exports are defined as too small, so the home sales are reduced accordingly. The home sales are reduced as opposed to increasing the exports so to ensure that the exports are not increased above the total sales producing negative home sales values, i.e. ProdCom estimates provide constraining totals. An example showing the workings of the new home sales method is available on request.

By introducing this method the following properties hold true:

- The home sales and exports are proportional to the ProdCom sales for the product within its industry.
- There are no negative home sales values.
- MPI data provides reliable estimates of export proportions by industry, which are coherent with ProdCom data because both

surveys are based on the same sampling frame with the same allocation of contributors to industry.

Comparison of results at the higher level using the old and new home sales method

To determine the impact of the change in method on results, a comparison has been made of index values produced with index weights derived on both the old and new home sales. The analysis undertaken focuses on the difference in the method of the home sales and does not include the change due to the base year being updated. All the data calculated is provisional and uses the new base year 2000=100 to calculate weights and indices but with home sales estimates derived using both the old and new method.

Results

Weights

The change in the method to calculate the home sales figures will obviously introduce differences in index weights. These differences are due to the estimates of home sales or the proportion of this value within the total. The extent to which the weight change will impact on results will depend on the size of weights and index levels.

By introducing the new method to calculate home sales the overall total has increased by around 4 per cent. This was anticipated given that it was thought that the old method produced estimates of home sales that were biased downwards.

Analysis of differences in weights derived on the old and new basis found, as expected, larger changes for product (6-digit) level weights within industry level series. There were 73 product groups showing a difference of 10 per cent or more out of 1,277 product groups in total. At the next level the difference was less with 12 industries showing a change of 5 per cent or more out of a total number of 229 industries.
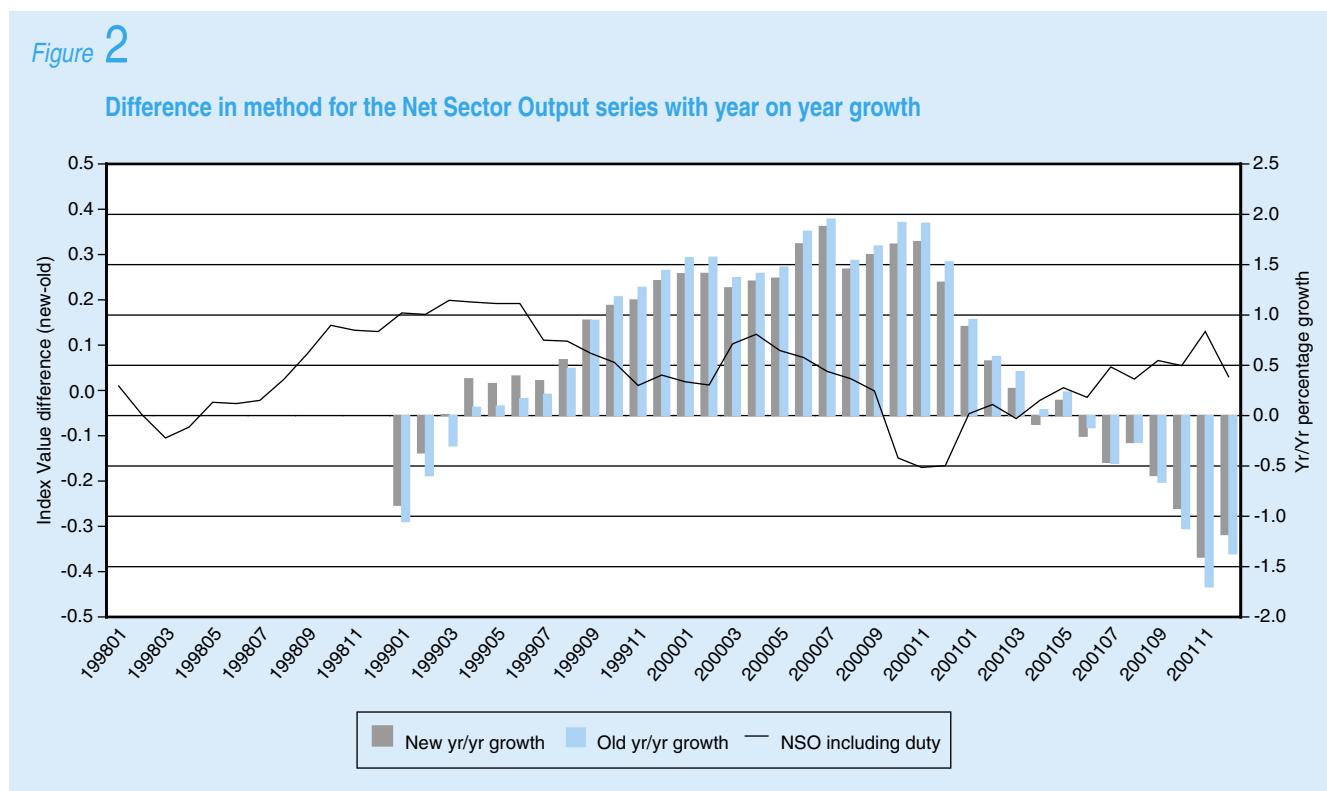
Investigation of these differences in weights identified the main reason being that under the old method when constraining the inconsistent data sets, exports were given too low a value.

Analysis of weight changes showed that, with relatively few exceptions, the new method provided an improved approach to the calculation of index weights.

Effect on the data series

What is of most concern and interest to users is the effect the introduction of the new home sales method will have on the data series published. This has been considered in isolation i.e. ignoring the effect of rebasing.

Figure 2 shows the provisional difference in NSO all-manufacturing index values due to the difference in methodologies – the difference shown being the series using the new home sales method weights less the series derived using the old.

Figure 2

Difference in method for the Net Sector Output series with year on year growth

Legend: New yr/yr growth | Old yr/yr growth | NSO including duty

It is worth noting that in calculating the NSO series it is only weights below divisional level that are affected. Under both approaches the weights of the division level series within the all-manufacturing index are consistent (and based on input-output data).

The provisional difference over time for the level is generally less than 0.2 index points with a few exceptions with the largest differences generally occurring towards the latter part of 1998 and the beginning of 1999. Differences in the year on year growth are on average less than 0.15 with the largest differences occurring towards the end of 2001.

Causes of differences in the NSO all-manufacturing series

As the weights for the division level series feeding into the high level are the same, then differences within the division level series will account for the all-manufacturing level differences. Figure 3 shows the provisional difference within the NSO series together with the contribution of the major component divisions to this difference.

Differences within the office machinery and computers component series are the main contributor to differences within the NSO series. Other components contribute to differences at the NSO level in specific periods, e.g. the alcoholic beverages component contributes to differences in specific months as do the chemical and radio, television and communication component.

As expected the differences between methods are greater at the lower levels most significantly within the office machinery and computer industry. With a small number of exceptions, provisional differences at division level are less than 2 index points. The reason for these differences is that there is now an improved approach which better reflects the products exported for these particular industries.

Weighting and structure of the import series

As part of the current rebasing project, improvements have been introduced in other areas. One such improvement has been to the import series. The import series like the output PPIs are constructed using price information from contributors and from published sources. A review of price quotes obtained from published sources has been undertaken to ensure that these are representative. This review resulted in the introduction of additional published source price quotes to improve coverage. The sample for both the import and export series has also been updated and increased again to improve the overall coverage and quality of the series produced.

The structure of the import series has also been reviewed with a view to introducing a consistent approach across all series. The structure of the series is now similar to that of PPI with detailed product level series produced based on a 4-digit Standard Industrial Trade Classification (SITC) classification. These product level series are in turn combined to produce 3-digit and 2-digit level SITC series

Figure 3

Contribution of specific component divisions to differences with the all-manufacturing Net Sector Output (including duty)
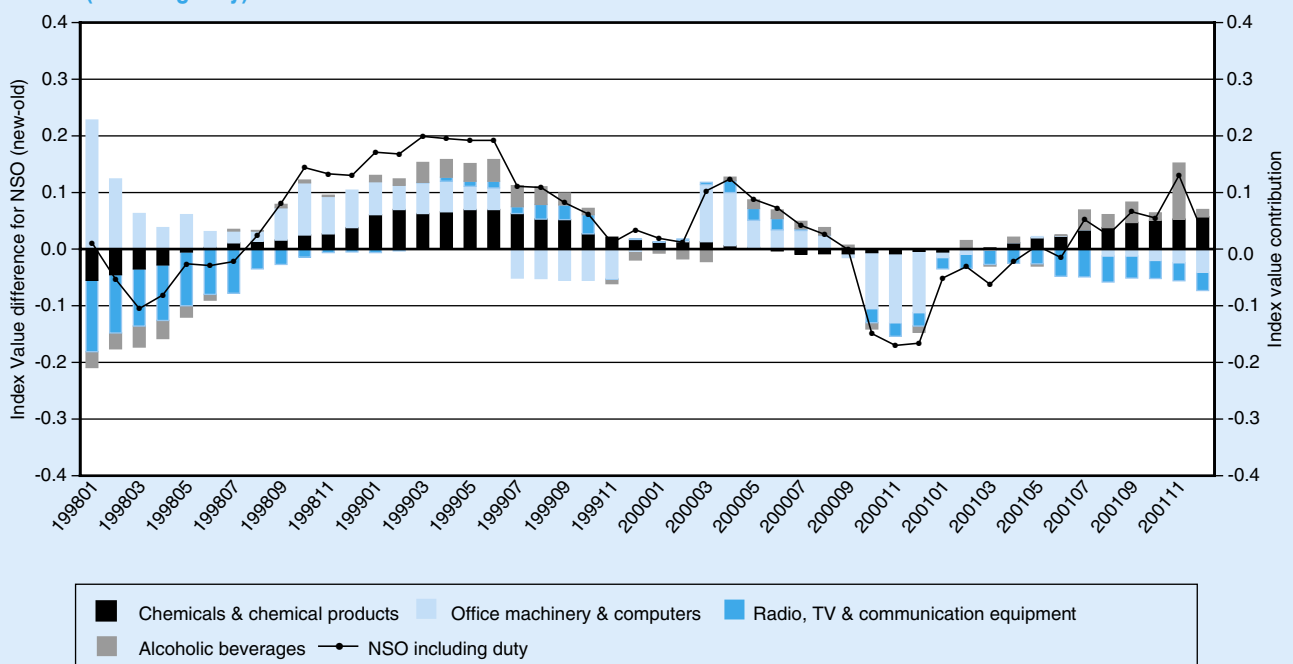
and all-import indices, with index weights based on relative import values of component series. There has also been a slight change to the approach used to weight quotes from businesses to form the product level series. Weights are now based on relative 2000 product import sales values of businesses included in the sample.

The input series review

The input series measures the movement of prices of materials/ products purchased by manufacturers that are required as inputs to the manufacturing process – both the raw material intermediate products that feed into the manufacturing process and materials required for the day-to-day running of the business. Input indices are constructed on both a gross and net sector basis. The net sector input indices are published at the all-manufacturing level only. This series excludes all inter-sector transactions, that is purchases from UK manufacturers are excluded from the gross purchase values before weights are calculated.

The structure of the indices is defined using input-output domestic and import matrices which provide information on the products purchased by industry groups from both the domestic and import markets. Due to cost and burden constraints it is not possible to collect input price data directly so gross sector output and import prices are used as proxies to construct the series. The review of the structure of the input series has considered which output PPIs and import indices should be used as components of the series. The review has considered:

- An appropriate structure for the series;
- Availability of component indices to be used in constructing the series.

In reviewing the structure of the series component, indices have been limited to raw materials and intermediate products – capital products have been excluded from the series. Since the last rebasing exercise the coverage of the import series has been improved. Suitable import series introduced since the last rebasing exercise have also been included as components of the input series. Whilst coverage of the import series has been increased there are still areas not covered by the current sample. In these cases proxy series have been used to represent the missing components. These proxy series are based on higher level import series.

A benefit of this approach is that this new structure allows for replacement of the proxy series with the appropriate import series, should coverage of the import series be expanded following completion of the rebasing exercise.

The impact of these changes were generally found to be less than 0.2 index points across all series and periods analysed.

Conclusions

This article has given details of the methodological changes implemented as part of the 2000=100 rebasing exercise for PPI. In particular concentrating on the changes made to the calculation of home sales data used to weight the output PPIs and identifying the effect this change has had on the high level series. The main points of conclusion are:

- The new method to calculate home sales are more robust compared to the previous approach.
- The differences due to the change in the method were small at the higher level with larger differences occurring at the more detailed level.
- Other methodological changes have been introduced during the rebasing programme. These are a review of how the input series is calculated and structured, and the structure and weighting of the import series.

These changes are to be implemented as part of the 2000=100 rebasing exercise. The analysis concentrates on the difference due to the methodological changes but the results of rebasing will also incorporate the change due to the update of the weights. The first results to be published on this basis will be available from October 2003. If you have any comments or if you would like further information with regard this article, please contact Jon Gough at the address given at the beginning of the article.

Reference

Morris L and Birch T (2001) Introducing a New Estimator for the Producer Price Index. *Economic Trends* No. 573, pp 63–71.