Sample design for Producer Price Indices

Paul A. Smith, University of Southampton, UK
Jillian Delaney, Central Statistics Office, Ireland

Abstract

The collection of price information is complicated because of the structure of prices charged by businesses. It is only practical to sample businesses at the first stage, rather than sampling prices directly, because this is the only frame information available. This leads to a multistage design, with the additional challenge that knowledge of which products an individual business makes is also needed, so that the sample size of price quotes can be controlled at the product level. There are different ways to operationalise such a sampling procedure, depending on whether there is already a source which contains some information on the products originating from a business.

In this paper we review the development and implementation of the UK's PPI sampling strategy, which changed to a probabilistic design in the 1990s, and document the challenges in setting up a path to develop the CSO's Wholesale Price Index to follow a similar probabilistic approach. The development of a probabilistic sample will be a long-term project, because of challenges around the availability of product-level data, and the response rates to business surveys in Ireland. The results provide an interesting case study of implementing a business price survey in a country with a smaller economy, notably including small numbers of large businesses concentrated in specific industries and therefore with production of particular products.

1. Introduction

Sample design for business surveys is challenging because of the skewed distribution of business sizes and the consequent need for substantial differences in inclusion probabilities, including complete enumeration of all the largest businesses. But the sample design process is also supported by good frame and auxiliary information, which means that the designs can be optimised to the characteristics of the business survey populations (Smith & Yung 2023). There is inevitably a timing mismatch between the frame information and the implementation of a design, since the design process itself requires some time to develop and implement, and the sample is generally deployed over a period. The lags have become smaller with the digitisation of administrative processes and increase in their use, as the mechanisms for data transfers to national statistical offices have improved.

The commonest approach to sampling for business surveys is targeted at outcome variables which are typically well correlated with business size (typically turnover, employment and similar variables). Some outcomes are less clearly correlated with size, such as research & development, or environmental protection expenditure, and then the design process may be more involved and may require identification of businesses that are more likely to have these kinds of activities. Here we focus on a further kind of survey, those for business prices. Price indices are constructed from two key pieces of information – weights indicating the sales of given products, and the prices of the products themselves. These are not needed with the same frequency, as the weights are updated only periodically, either at a rebasing exercise or annually as part of a chainlinking process for the indices. The prices, however, are needed for each instance, as they are the part which varies and these changes are what the index aims to measure.

The USA was the first country to introduce a probabilistic approach to sampling in its producer price index (PPI), as part of a wider strategy to have probability sampling in all its price indices, and essentially the whole selection was probabilistic from 1978. The UK also moved to a probabilistic design, for selection of businesses and price quotes, but not for the selection of the items to be priced within product categories, in the late 1990s. An overview of the approaches in these two countries is

given in Delaney & Smith (2021). Other countries have followed the UK and introduced probability sampling approaches for businesses and price quotes, including Sweden and the Netherlands, though many countries still operate cut-off type designs for business price indices. In this paper we review the development of the sample design for the PPI in the UK, and then consider how such an approach might be implemented in Ireland where less data is currently available, and a gradual strengthening of the sample is the most practical approach.

2. Sample design for the PPI in the UK

The UK was able to move to a probabilistic design when the PRODCOM product survey was introduced, as it had a probabilistic design which could be used as the first phase a two-phase design for collecting prices. (Before that the product surveys were based on cut-off designs which did not lend themselves to a probabilistic framework). PRODCOM provides the sales by detailed product categories, and these are used to form business-product pairs. Each such business-product pair has an associated sales value (from the business's response to the PRODCOM survey), and when multiplied by the sampling weights these form the index weights.

A second phase is then needed to select some of these business-product pairs, each of which results in a business being approached to give (in principle) one price quote. Using a single quote means that the effect of the clustering of prices within businesses has least impact on the accuracy of the resulting index. This second phase also generates a sampling weight to compensate for the differential sampling. The form of the estimator of the producer price index is therefore

$$ppi_j = \frac{\sum_{\substack{ij \in s_h \\ i \in s_g}} w_g w_h s_{ij} pr_{ij}}{\sum_{\substack{ij \in s_h \\ i \in s_g}} w_g w_h s_{ij}} \qquad (1)$$

where $w_g = N_g / n_g$ is the weight for business *i* in stratum *g* of the first-phase design (ie PRODCOM), $w_h = N_h / n_h$ is the weight for business-product pair *ij* in stratum *h* of the second phase design (ie PPI), s_{ij} is the sales of product *j* by business *i* in the PRODCOM collection, and pr_{ij} is the price relative for product *j* in business *i*. Note that the information from PRODCOM is for year *t*-2 when the PPI information is for year *t*, so there is some potential for mismatches in the information.

In principle, model-assisted weights (such as ratio or calibration weights) could be used to replace the expansion weights, but both PRODCOM and PPI sampling use relatively small sample sizes within strata, so the expansion weights are more robust (less susceptible to unusual samples within strata).

Equation (1) is used to estimate the lowest-level indices, which are at the 6-digit level of CPC-COM, and these can then be aggregated to 4-digit indices (the publication level) using weights, also estimated from PRODCOM. These can then be further aggregated to a whole-economy PPI in the same way.

We can approximate the sampling variance of (1) using Taylor linearisation (Purdon et al. 1998). This leads to a complex expression with three terms, which represent the variance due to the first phase of sampling from Prodcom; the variance due to the second phase of sampling of business-product pairs for the PPI; and a variance due to the sampling of products within a business-product pair. This last is still a purposive process in the UK, although in the USA this is also probability-based, with successive listing and pps selection until a defined product (with all its characteristics) is selected. The 6-digit product variances can be weighted together to give the 4-digit variances; because the selections of the products are independent in different 6-digit groups, the second and third terms in the variance can simply be added (with squares of the appropriate weights). The first term, however, involves the selection of different products from the same business, which means that there is some correlation. Fortunately the variability at this stage is small relative to the other two components of

variance (Purdon et al. 1998), so a simpler approximation of the variance can be used with only a small effect.

In order to estimate the components of the variance, several inputs are required. The third term, for the sampling of products within a business-product pair (ie sampling from the products within a certain classification in the selected business) requires knowledge of the number of different products Q_{ij} from which a sample of q_{ij} products will be selected. This is information that is unlikely to be available; for the optimisation of the design in the UK, a value of 5 was assumed for this component for all business-product pairs. The associated finite population correction was assumed to be negligible, which results in an overestimation of the variance, so that values are conservative.

The other components of variance were originally estimated under a model relating the PPI and PRODCOM variability, because there was not enough linked information to estimate them together; this could potentially be revisited.

The variances have two roles – one is to provide information on the sampling variability of the PPI indices so that users can make an assessment of the quality of the outputs. The other is as an input to the design process, so that the accuracy of the resulting indices can be optimised for a given level of inputs. The PRODCOM element of the sampling is designed for the purpose of producing PRODCOM outputs, and its stratification is regarded as fixed. The stratification of the PPI selection can however be decided. An examination of the formula for the variance shows that the two terms which are affected by the PPI stratification both depend on equations involving s_{ij}^2, so we can reason that stratification on sales values s_{ij} will be appropriate. Using a simple cumulative aggregate size approach which ensures that the total sales values within each stratum are approximately equal is argued by Purdon et al. (1998) to be close to optimal; however, it would equally be possible to use the Dalenius-Hodges allocation procedure for more detail.

The number of sales strata could be varied, but in the UK implementation is fixed at 3, with the largest stratum completely enumerated. The stratum boundaries could also be varied for each 6-digit product, but there is insufficient data to support estimation of the boundaries at this level, and the additional complexity for systems is probably not worthwhile, so common boundaries are set for all 6-digit products within a 4-digit product group.

Finally an optimisation of the sample is undertaken, based on a search for the variance which can be achieved for a certain sample size. The equation for this contains a term of the form $V - const$ where V is the target variance and $const$ is a sum of constant terms which are not affected by the optimisation. If $V > const$, then the target variance can be achieved. Additional constraints are needed to ensure that there are enough price quotes in each product, and this may involve some iteration where certain strata are given a fixed sample size and taken out of the optimisation.

As with many sample optimisation problems, there is a chance that the allocated number of prices to be collected exceeds the number of business-product pairs available at the second stage. In this case there are two options:

- use only the number of pairs available, and reallocate the unused sample elsewhere
- sample more than one product within each pair

The latter is expected to be less useful than sampling a different business, because of the correlation in prices within the same business, but may be better than having a small sample. In the UK implementation of the design, a constant number, which could be >1, was chosen from each pair.

2.1. Calculating sampling variances

The process described above was based on a Taylor linearisation of the index formula (1), and contained many approximations. Some doubt was expressed about their efficacy, particularly in the case of small sample sizes which result from this design process. Therefore a new method for calculating estimates of the variance was introduced based on a model incorporating a series of

assumptions by Wood *et al*. (2008). Some of the assumptions are investigated, and it is interesting to see that the contemporaneous correlation between price movements is non-negligible (ie there is clustering in price movements). The model incorporates correlations between different price changes within the same product group for the same business, and also between price changes in different product groups for the same business. The resulting variance estimators are rather complex, but have since been used as the basis for calculating sampling errors for the PPI in the UK, with the most recent estimates published in 2019. A recent review by the Office for Statistics Regulation (2023) has recommended that the regular calculation of these measures is reinstated.

2.2. Unanswered questions

The sampling and variance estimation calculations have been developed for PPIs in the UK, but are relatively complex, and the effects of the assumptions in the different methods are not easy to assess. Some comparative analysis was undertaken when the new variance estimation method was introduced. The PPI sample size in the UK has been reduced since the initial design work was undertaken, and I believe that the optimisation was undertaken with the new model-based variances, though further investigation to confirm this is needed.

Several methodological innovations could be investigated. In particular, the use of stratification at the second stage of the sampling is practical and robust to changes in the sales in different product categories caused by the timing mismatch between the PRODCOM data used in the design and the price collection. But for the same reason it is not well optimised, particularly for small sample sizes. Pps samples are known to be more efficient when small sample sizes are used; some examination of the robustness is needed, and this could initially be based on the product sales series to examine their volatility.

One of the reasons for using a probability sampling design is to mitigate the risk that price changes are correlated with the size of businesses. Since the introduction of these designs, substantial amounts of data have been collected which could be used to investigate whether this is a real phenomenon. The correlations in pricing behaviour within a business (both within and across product groups) could also be revisited.

3. Sample design for the WPI in Ireland

The Wholesale Price Index (WPI) is the Irish release that includes industrial Producer Prices. It has not been redesigned for some time, and is in need of development to ensure that it is measuring prices as accurately as possible within the available resources. The Irish industrial PPI is currently based on a sample of units taken from the Census of Industrial Production, and relies on a questionnaire to gather product data as the basis for sampling price information. Although completion of the questionnaire is compulsory under Irish law, the response rate is rather low (Delaney & Smith 2021), and any businesses not responding cannot be included in the price collection, so there is a risk that the nonresponse bias will be significant.

Delaney & Smith (2021) give an overview of the current design of the Irish PPI, and some indications of how the design could be modernised. Since then we have investigated further the landscape of surveys. Following a probability-based design would be a good, defensible methodological strategy, but there are some features of the Irish statistical system which mean that it is not possible to change to a two-phase probability-based design of the type which we have just outlined from the UK's implementation. The Census of Industrial Production (CIP) is an effective survey for gathering information on local units in Ireland with 3 or more employees, but does not include any product sales information (even at a fairly aggregate level) among its variables. It is therefore not usable directly as the first phase. PRODCOM also runs in Ireland, and this looks like an obvious choice as a first phased because it does collect product information. However, it has lower priority among the limited resources available for nonresponse follow-up, so that, although response is in principle compulsory,

the actual response rate is currently around 35%, which is not high enough to support a two-phase design.

Instead we propose a hybrid approach, which uses PRODCOM responses where they are available as the basis for sampling, giving one part of the design, and following the two-phase design principle. These units would be issued a canvass form requesting details of a sample of product transactions in the sampled PRODCOM codes. For units which do not respond to PRODCOM, use a canvass questionnaire to gather product information, as now, where the respondents are asked to choose their main products. In both cases, the data from the canvass questionnaires are used as the basis of price collection, giving the second part of the design.

More details about this proposed approach:

- the sample should initially be formed of two components:
  (a) sampling of business-product pairs from a frame constructed from responses to a previous PRODCOM. Products should be stratified by the industry that makes them to maintain the current approach of collecting prices for products of an industry, (suitable for deflation) rather than a homogeneous set of products;
  (b) sampling of businesses (currently LUs, but moving to KAUs in 2024 in line with the European Business Regulation) from a previous CIP, excluding any businesses which already appear in the PRODCOM frame.

  For part (a) PRODCOM responses provide details of product groups and their sales allowing the CSO to ensure that products in these groups are included in the PPI sample. Part (b) would continue to use a canvass questionnaire to gather information on products; the CSO is working on better enforcement of the compulsory response to this questionnaire.
- there is an extra step in allocation to decide how much of the sample should be used in part (a) and how much in part (b). Ideally this should be based on variability of prices, but a proportional allocation by number of businesses may be a straightforward initial approximation.
- use πps selection for business-products pairs in part (a), because it deals most efficiently with relatively small sample sizes in populations of units with diverse sizes. The standard application of πps selection results in some large units which are certain to be selected in the sample, so this divides the sample into CE and sampled parts (see for example Blijenberg (2007, section 2) for a description of how this works). We note that the UK uses a size stratification of business-product pairs, and thereby avoids πps selection, but the extra stratification adds complexity to the design as the stratum boundaries need to be maintained, and is less efficient in its use of a relatively small sample. The inverse of the sampling probability should be used in the calculation of the weight to feed into the index calculation, so that the proper representation of the population is maintained.
- selection for part (b) is simpler because it covers only the businesses, so can probably use stratified simple random sampling. Since the sales value by Prodcom code is not available from the canvass, πps selection is not practicable. The inverse of the sampling probability should be used in the calculation of the weight to feed into the index calculation, so that the proper representation of the population is maintained. A question of how to deal with non-response to the canvass remains; one option is to treat the achieved sample as if it were the selected sample, but since the response rate to the canvass is rather low, this risks a potentially large nonresponse bias. There may, however, not be any better information with which to make a nonresponse adjustment.
- the part (a) sample should be updated each year using a suitable rotation scheme. Annual updating will ensure that the sample (particularly the weights) keeps pace with changes in the economy. The CE part of a πps selection will ensure that the largest businesses are continuously included. A rotation scheme is needed for smaller units so that they can be retained in the sample for >1 year, to maximise the benefit once they have been “trained” in

how to respond. Nevertheless, rotation should allow such units to drop out after a fixed maximum time, in order to spread the burden of responding. (Larger units would be expected to remain in the sample 'forever').

- The connection to Prodcom should mean that a weight is always available in part (a), so that minimal weights are not needed. In part (b) we would like to avoid minimal weights, which mean that CSO puts in a lot of effort to collect price quotes which have only a small effect on the index. If product sales are not available, it is not clear how this information can be effectively gathered or approximated, and the best approach has not yet been resolved.
- regular assessment of the accuracy of the PPI by calculation of suitable sampling errors can then follow using one of the approaches from Purdon et al. (1998) or Wood et al. (2008).

Once this approach has been set up, the idea would be to improve the response rate to PRODCOM, as a basis for improving the second-phase sampling for the PPI as well as improving the PRODCOM outputs. There would be an important advantage in collecting product transactions for specific Prodcom codes in part (a). This would allow CSO to control the sample to ensure that important product groups are included rather than taking what businesses indicate are their main products. Eventually, when the response rate to PRODCOM is sufficiently high, the element from the CIP can be dropped.

4. Conclusions

The conditions in the Irish statistical system make it difficult to transition directly from the existing PPI sample to one based on a two-phase probabilistic design. But there is a way to introduce a hybrid design which makes use of the PRODCOM information for one part, and continues with the existing procedures in a second part. Overtime the size of the PRODCOM-supported part should grow, and eventually the whole operation will be based on a probabilistic design process.

References

DELANEY, J. & SMITH, P.A. (2021) Sampling for business price indices, with applications in Ireland. *European Establishment Surveys Workshop EESW21, on-line.* Available from: https://drive.google.com/file/d/1nBkF0p158CGILL7Fbm_ISfQV_rJVS4Rg/view

PURDON, S, COPE, I. & DAVIES, G. (1998) *Obtaining an optimal allocation and weighting system for the United Kingdom Producer Price Indices*. Report to the Office for National Statistics.

SMITH, P.A. & YUNG, W. (2023) Introduction to sampling and estimation for business surveys. Chapter 27 (pp615-636) in Snijkers, G., Bavdaž, M., Bender, S., Jones, J., MacFeely, S., Sakshaug, J.W., Thompson, K.J. & van Delden, A. (eds) *Advances in Business Statistics, Methods and Data Collection*. Wiley: Hoboken.

OFFICE FOR STATISTICS REGULATION (2023) Spotlight on Quality - Producer Price Indices. Assessment Report 378. UK Statistics Authority. Available from https://osr.statisticsauthority.gov.uk/publication/spotlight-on-quality-producer-price-indices/

WOOD, J., ŠOVA, M.G. & BUCKNALL, R.D. (2008) The Method used to estimate Standard Errors for UK Producer Price Indices. *Survey Methodology Bulletin* **61** 62-80.