Quality adjustment in the New Zealand Consumers Price Index

Current practice and future directions

Note – this is a chapter (unedited version) for
The NZ CPI at 100: History and Interpretation
to be published by Victoria University Press, Wellington in 2014

Frances Krsinich

Senior Researcher, Prices, Statistics New Zealand

Email: frances.krsinich@stats.govt.nz

Abstract: We explain the range of different approaches for dealing with quality change currently used in New Zealand Consumers Price Index. Since 2001 these have included a hedonic regression approach to estimating the price movements of used cars. A fixed effects index was used to benchmark the current matched-sample approach to quality-adjusting housing rental price movements. New data sources, in particular scanner data for consumer electronics and supermarket data, and web-scraped online retail data, present new challenges for quality-adjustment. We summarise the recent methodological research contributed to by Statistics New Zealand in this area. Exploiting the longitudinal nature of the data in the hedonic framework appears to have significant potential for quality adjustment.

1 Introduction

A key issue when estimating price indexes is ensuring that only 'like-for-like' price change is reflected in the index. Products are constantly evolving over time, with associated changes in quality. If this quality is not adjusted for appropriately, price indexes will be correspondingly biased. For example, if a can of beans reduces in size but its price remains the same, this should contribute an increase to the price index.

The Boskin Commission report on the CPI in the United States of America (Boskin, Dulberger, Gordon, Griliches, & Jorgenson, 1996) highlighted the importance of appropriately adjusting for quality change to ensuring a fit-for-purpose CPI.

Historically, quality-adjustment in the NZ CPI has been dealt with by the maintenance of a fixed and representative basket of goods and services. Section 3 summarises the different approaches to quality-adjustment when goods and services initially in the fixed basket are no longer available for pricing.

Regular updating of the basket – in terms of the products priced, the outlets they are priced at, along with their associated expenditure weights, keeps the basket representative of what New Zealand consumers are buying.

New data sources, such as scanner data and online data, give us significant opportunities to improve the quality and cost-effectiveness of estimating price change (Statistics New Zealand, 2013). But quality adjustment becomes more challenging when using this kind of data. Traditional price index formulae don't work with such big datasets because of the high level of turnover in products sold, and the volatility in prices and quantities due to sales.

'Hedonic regression' is a common approach to deriving quality-adjusted price change, by statistically modelling prices against time and the price-determining characteristics of the products. The price index is then derived from the parameters estimated for time. Weighting these models according to quantities ensure that the measures of price change reflect the expenditure shares appropriately.

The hedonic approach is an obvious one for quality-adjusting price indexes from thsee new data sources. However, information on the product characteristics are not always available in the data. Our research has shown that there is potential for exploiting the longitudinal nature of price data to counter this lack of characteristics information and produce appropriately quality-adjusted price indexes.

2. Historical view on dealing with quality change

The following selected extracts from give some insight into the evolving focus on quality adjustment in New Zealand price measurement historically.

The 1912 Royal Commission on the Cost of Living emphasized the importance of pricing constant quality products across time, noted that quality is generally improving, but concluded that this could not be adjusted for 'in the price statistics themselves':

> In dealing with the material accumulated regarding prices, the Commission was alive to most of the leading principles to be observed in combining and interpreting them, viz:-
>
> …That the things priced should be the same not only in kind but also in quality. This is a matter impossible to allow for in the price statistics themselves, as they cannot be manipulated to allow for improvement in quality. But it is important to note that most of the articles with which the inquiry has been concerned have improved in quality during the last twenty years—some of them very considerably; that is to say, not only has there been an improvement in quality in each grade of the commodity, but in many cases new and higher grades of the same article have appeared. If ordinary white bread, e.g., improves in quality and the money price remains the same there is a real fall in price, but there is no plain and convenient method of representing statistically such a fall. All that can be done is to call attention to the fact that the quality has changed. (New Zealand Government, 1912 - p12)

The 1948 report of the Index Committee discusses the requirement for clear specifications to ensure identical qualities of goods are being priced:

> A further point of considerable importance is the effect on family expenditure of changes in the quality of goods. It is, of course, essential that, where possible, the prices reported should relate to identical qualities of goods at each period of reporting. For this reason the Committee stresses the importance of careful specification of each item to be priced; and the Government Statistician has drawn up price-collection forms after close consultation with expert opinion as to specifications for each item.
>
> Even with such precautions, difficulties in ensuring quotations for comparable qualities at each price collection are not easily overcome; and supervision of this aspect of the price-collecting programme will form an important part of the work of the price-collectors.
>
> …While it is difficult to lay down any hard-and-fast rules regarding the reflection of changes in quality in a retail price index, the Committee considers that changes of importance should be taken into account, since a marked depreciation or appreciation in the quality of an article priced is equivalent to a price change and should be reflected in the index. In some countries samples of the goods to be priced are kept by the price-collectors; and; for the more difficult items, this practice is recommended. (NZ Government, 1948 - p9)

And recommendation 21 of the 1978 report of the Consumers Price Index Revision Advisory Committee stresses the importance of ‘quality control of price data’:

> That in order to maintain the statistical standards of the index the work of quality control of price data be pursued to the maximum degree possible and in line with accepted international practice.(NZ Government - p28)

In 1985 the committee supported the intention of the then Department of Statistics to introduce a price-relative method, noting its ability to better deal with the quality adjustments required by substitutions for obsolete commodities.

> The price-relative method avoids the problem of incorrect implicit variable outlet price weighting, because absolute prices are not averaged. It also offers greater flexibility in handling on-going practical problems associated with actual compilation of the index such as quality adjustments resulting from the substitution of obsolete commodities in individual outlets. The committee supports on these technical grounds the department's intention to introduce the price-relatives method of calculating the within-town price-change averages as soon as it is practicable to do so. (NZ Government, 1985 – p17)

And the report of the committee also notes that quality-control is a challenging issue in practice:

> 12.9 Only 'pure' price changes should be incorporated into the index calculation. Excluded are price movements, or contributions to price movements, which arise from variations in quantities or qualities of surveyed commodities. This is not a trivial task as change is endemic in the market place. The Department of Statistics has therefore had to develop techniques to not only detect change, but also to suitably quantify its effects on price. In general, this area of work is known as 'quality control'.
>
> 12.10 The quality of a commodity is determined by the quantity and standard of service or satisfaction it will provide to the customer. At a practical level it is, in most cases, very difficult to define the quality of a service with any objectivity, normally making it necessary to assume this remains unchanged over the life of the index.
>
> 12.11 In the case of goods, any differences in the physical characteristics (for example size, style, operating characteristics, etc.), and conditions of sale, such as associated customer services, may be considered quality differences in terms of index concepts. While the former can usually be objectively assessed the latter usually necessitate some subjective appraisal. Other criteria for assessing quality differences, such as functional efficiency and perceived value, are also dependent on individual interpretation. Quality control procedures are developed to minimise the element of subjectivity. (NZ Government, 1985 – p41)

The report of the 1991 revision advisory committee defines more clearly what is meant by ‘quality’ in the CPI:

> The fundamental conceptual principle of quality control in the CPI is that the quality of a commodity is determined by the quantity and standard of service or satisfaction it will provide to the consumer. This definition that quality is determined by perceived value to the consumer provides a different assessment of quality to one using functional efficiency or the manufacturer's specification as a criterion.
>
> The "quality assessor" must calculate, or else make judgements as to, the difference in perceived value to the consumer between the original and replacement commodities. Quality control procedures are developed to try and minimise the element of subjectivity involved.
>
> In the case of goods, differences in quantity, componentry, size, style, packaging, functional or operating characteristics (e.g. durability, capacity and speed) and associated customer service at the point of sale are all examples of quality differences. For basic consumer goods these can be assessed with a reasonable degree of objectivity. However for more complex durable goods this is not so simple. A new motorcycle model may feature a bigger fuel tank and a more powerful engine, yet have poorer handling than the model it replaces. The challenge is to make a combined assessment of all the diverse quality changes. (NZ Government, 1991 - pp45-46)

And the 1991 committee expressed its confidence in the quality control practice of the time:

> The Committee noted with satisfaction that by international standards, the quality control techniques and procedures employed in the New Zealand CPI are particularly robust and expressed its confidence that the CPI is as good a measure of price level change of a basket of goods and services of constant quality as is technically and practicably possible. Furthermore the Committee considered that the Department of Statistics should do more to publicise its efforts with respect to quality control procedures in the interest of promoting confidence in the reliability of the CPI. (NZ Government, 1991 - p27)

3. Current quality adjustment practice in the NZ CPI

Currently, Statistics NZ uses quality adjustment methods consistent with international best practice outlined in the International Labour Organisation's (ILO) resolution on CPIs (International Labour Organisation, 2003). These practices ensure the CPI is fit for purpose by addressing both the representativeness of the CPI basket and the need to maintain a fixed quality. They include:

- updating the CPI basket and weights once every three years
- reviewing product and retail outlet samples regularly
- making quality adjustments when products tracked in the index have changed characteristics
- using hedonic modelling to quality adjust the price movements of used cars

The characteristics of the products being priced are monitored to ensure that any relevant differences can be excluded from the estimated price change. These changes can be in the size, performance, or functionality of the product. Quality assessments put a monetary value on the change in characteristics between the old and new product, as perceived by the consumer. Several different techniques or methods are used, depending on the type of good or service being tracked.

3.1 Product or outlet unavailability

At times, products scheduled for pricing are not available – they may be temporarily out of stock or permanently removed from sale. If a product is temporarily out of stock, the price from the previous period is carried forward, and alternative products are identified for possible future use.

Usually, if the product is still unavailable at the next pricing period, it will be replaced with one of the products previously identified as a suitable replacement (or if necessary, the outlet is replaced). If a class of item is removed from sale from an outlet, then it will be tracked in the future, where possible, in a similar outlet.

3.2 Cheapest available specifications

For many products in the CPI basket, a specific product of a specific brand is tracked consistently. For others, the cheapest available option is tracked, regardless of brand. This is done where products are judged to have little quality difference across different brands, and for which consumers are unlikely to show much brand loyalty. For these products, tracking the price of one particular brand would not represent consumers' buying patterns as well as tracking the 'cheapest available' specification would.

Examples of products where the cheapest available option is priced are frozen peas, white bread, white flour, standard homogenised milk, cheddar cheese, standard eggs, butter, and sugar.

All fresh fruit and vegetables are priced on the basis of cheapest available, as long as the variety chosen is of suitable quality for most uses. This approach also allows for a consistent pricing pattern in products where specific brands or varieties are not reliably available at all sampled retailers.

3.3 Changing pack sizes and quantity specials

A common example of explicit quality adjustment is related to a change in pack size. For example, tea bags usually sold in boxes of 100 bags may come with 10 percent extra due to a promotion run by the distributor. In this case, consumers receive the benefit of an extra 10 bags, so the recorded price is adjusted to reflect the value of the extra tea bags to the consumer.

Similarly, 'quantity specials' are also taken into account. For example, if a loaf of white bread was $2.00 in March 2011, and three loaves of white bread were $4.80 in April 2011 (the single price remained $2.00), then the price per loaf has decreased from $2.00 to $1.60, so a 20 percent fall in price would be shown. Such quantity specials are reflected only where they are considered to be representative of the quantities likely to be purchased by households.

3.4 Option costs

For some products, quality adjustments are based on the market value of optional features, or a proportion of the market value of optional features. This approach is currently taken for desktop computers.

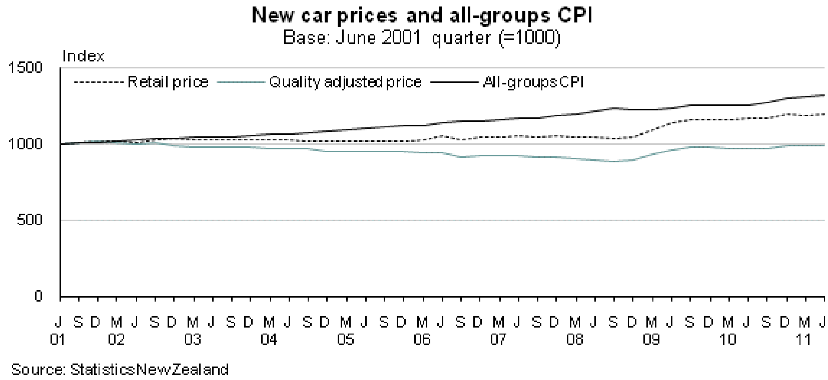
In the case of new cars, when distributors report changes to the sample models, we ask for the ‘perceived’ dollar value of these changes to customers. To ensure the adjustments are consistent, these are checked against records of previous adjustments. When there are changes to the engine, the value of the quality change is estimated on the basis of maximum power and torque. The values of all changes between the two models are combined.

Most quality adjustments to new car prices are made to remove the effect of improved or additional features, which increase the quality of the vehicle. In these cases, the value of the changes is removed from the retail price of the updated model to generate the like-for-like price.

If, on the other hand, the quality adjustment is due to a removed or diminished feature, the value of the change is added to the retail price of the updated model. In the 10 years from 2001 to 2011, most adjustments for lower quality were due to reductions in engine power or torque.

Figure 1 shows that over the 10-year period from the June 2001 quarter to the June 2011 quarter, the index for new cars based on retail prices increased 19.1 percent. The quality adjusted price series, which is used to calculate the CPI, decreased 1.5 percent over the same time period. The large difference between these two price series – which implies an annual average increase in quality of 1.9 percent – shows the importance of quality adjustments in the CPI. See Statistics NZ (2011a) for more detail.

Figure 1. Retail price and quality-adjusted price of new cars 2001-2011



3.5 Hedonic regression

Statistics NZ runs a quarterly survey of used car dealers, which collects prices and characteristics of approximately 3,500 cars sold. A statistical technique called hedonic regression is used to estimate the price index for used cars from this data.

A used car can be seen as comprising a bundle of price-determining characteristics. Once these characteristics are identified and measured, the hedonic model can be interpreted as breaking down the car's price into the implicit prices and quantities of each characteristic. The price index can then be derived from the estimated price over time, after controlling for the changing quality composition of the cars being sold from quarter to quarter.

The hedonic method for used cars was implemented in the September 2001 quarter and updated in the June 2011 quarter (Statistics NZ, 2011b). Section 4 describes hedonic regression and its application to used cars in more detail.

Hedonic regression was also briefly used to measure price change for whiteware in the early 2000s but this was discontinued as the approach was resourceintensive and resulted in index movements that were very close to those from the existing method.

3.6 Implicit quality adjustments

Section 3.1 above described what happens when products are unavailable. If a product is temporarily unavailable for two consecutive time periods, then a suitable replacement product, identified at the first instance of unavailability, will be used as a permanent replacement. In some instances it is assumed that any difference in price between the original and the replacement products reflects a difference in quality. This technique is called the 'overlap' method.

Statistics NZ uses the 'comparable replacement' method when the replacement product is judged to be very similar in quality to the old product, such as a newer model with only small, superficial changes. In this case, any change in retail price between the old and new models is shown in the price index.

When the replacement is judged to be of different quality to the old product, we in some cases infer the price movement from products that are directly comparable from within the same geographic region as the product being replaced. This is called the 'class mean imputation' method.

For some products, quality is implicitly controlled by calculating price change based only on products that are available in consecutive time periods. This is a ‘matched-model’ or ‘matched-sample’ approach. Section 5 discusses the current use of a matched-sample for the estimation of quality-adjusted dwelling-rental price change, and its evaluation against a hedonic benchmark.

4 Used cars hedonic index

4.1 The used cars survey

The measurement of price change for used cars is particularly challenging. The concept of a ‘fixed basket’ is difficult, if not impossible, to operationalise in this context. For example, even if we could measure the price of the same used car at a later date, it will have a different quality due to being older and having travelled further. So a different approach is needed to measure like-for-like price change for used cars.

Each quarter, data on the sales of approximately 3,500 cars are collected from a sample of used car dealers. Price, year of manufacture, make and model, engine size (cc rating), and odometer reading are collected for each car sold.

The sample was designed initially to support the calculation of average prices within estimation cells based on combinations of make and model, cc-rating ranges, and age of car. To ensure robust estimation of averages, cells with too few observations were excluded, which resulted in only around 25 percent of the data being used. The average prices within estimation cell were weighted together then the change in used-car price was derived from these.

4.2 Introduction of hedonic estimation for used cars in 2001

In order to make more efficient use of the used cars survey data, a hedonic estimation was introduced in 2001. Price was modelled against quarter (i.e. time), region, make and model, age of car, cc rating, and odometer reading, as shown in equation (1):

$$P_c = \sum_k \beta_k C_{kc} + \sum_t \delta_t D_{ct} + \varepsilon_c \qquad (1)$$

Where:

P_c is the price of car c (Note that there is no t term as we can assume each individual car – at a given quality - is only sold once)

$D_{ct} = 1$ in the quarter t that car c is sold and 0 otherwise

C_{kc} is the kth characteristic of car c, where the k characteristics are:

- town of purchase (15 categories)
- make and model (47 categories)
- age (in years)
- size of engine (cc rating, eg 2300)
- odometer reading (in kilometres).

ε_c is the error term, distributed normally

The quality-adjusted price index is derived from the parameter estimated on time – i.e. the δ_t

In 2011 the hedonic model was improved, by

- modelling the log of price, which fits the data better
- incorporating a more detailed classification of make and model
- adding squared terms for age of car and cc rating
- adding an identifier for used car dealer to the model

The updated model is now:

$$\ln P_c = \sum_k \beta_k C_{kc} + \sum_t \delta_t D_{ct} + \varepsilon_c \qquad (2)$$

where the k characteristics are:

- town of purchase (15 categories)
- dealer (approximately 300)
- make and model (96 categories)
- age (in years) and age squared
- size of engine (cc rating, eg 2300) and cc rating squared
- odometer reading (in kilometres).

Hedonic indexes are calculated separately for each of five broad regions of New Zealand (Auckland, Wellington, Rest of North Island, Canterbury, and Rest of South Island) and these regional-level indexes are incorporated into the NZ CPI.

The NZ CPI is non-revisable so, to make the hedonic method operational, a rolling window of the latest eight quarters is used to estimate the hedonic model each quarter and the index movement for the most recent quarter is linked to the previous quarter's index number.

Statistics NZ (2011b) discusses the used cars hedonic regression method in more detail.

5 Hedonic benchmarking of the matched-sample estimation of the rental index

5.1 The residential rents survey

The residential rents survey uses an area-based probability sample of approximately 2,800 rental dwellings. Landlords of the selected dwellings are asked each quarter for rental information. The sample is updated every quarter with dwellings that enter the rental market within the selected areas. These are identified from data on bond [1] payments by new tenants, by reconciling dwellings with newly lodged bonds against the existing sample, within sampled geographic areas.

Average rents are estimated from the sample within strata based on number of bedrooms and five broad regions. For example, average rent is estimated for three-

[1] A bond is an initial payment made to the landlord, and lodged with the Ministry of Business, Innovation and Employment at the start of the tenancy, which is held until the end of the tenancy as a guarantee against damage to property, or non-payment of rent.

bedroom rental dwellings in the Auckland region. Movements in these averages are weighted together using regional population-based expenditure shares to create a national index. The expenditure shares are updated at three-yearly CPI reviews.

A feature of the sample is that each quarter approximately 6.6 percent is due to 'births' – i.e. dwellings new to the rental population, and around 3.2 percent are deaths[2]. This reflects the increasing size of the rental population in New Zealand, as a result of population growth and a shift away from home ownership.

5.2 Introduction of a matched sample in 2000

The rental survey was initially designed in 1998, and in 2000 it was modified by restricting the sample contributing to the price movement estimate, to those dwellings existing in both the previous and current quarters – that is, a 'matched sample'. This was done to ensure that differences in the composition of the sample due to newly rented dwellings and dwellings leaving the rental stock would not contaminate the estimation of price movements. This is a common approach to quality-adjustment in price indexes.

5.3 Hedonic benchmarking exercise

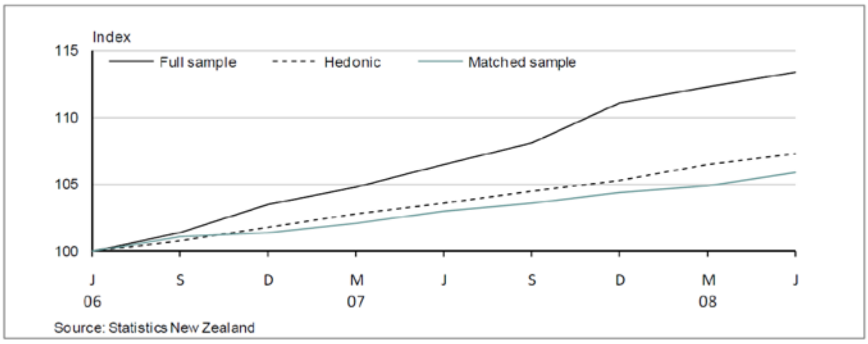
Smith (2008), of the Reserve Bank of New Zealand, speculated that the use of a matched sample might be biasing the index downwards by not reflecting like-for-like price change associated with the dwellings new to the rental market (and, to a lesser extent, those leaving the rental market). As rents tend to follow a 'stepped' movement, with rent increases more likely when tenancies change and (implicitly) when dwellings are newly rented, this is a valid concern.

Krsinich (2009) measured the significance of this potential bias by constructing a benchmark hedonic index from the rental data. This showed that any bias due to the matched-sample estimation is likely to be missing the implicit initial rent hikes associated with a newly rented dwelling was small.

Figure 2 shows the matched-sample index compared with the hedonic benchmark and to the index based directly on the full sample – i.e. based on the sample with no matching to ensure constant quality.

Figure 2. Comparison of matched- and full-sample indexes to the hedonic benchmark June 2006 to June 2008

[2] On average over the 2007 and 2008 years



Note: Base: June 2006 quarter (=100)

5.4 Fixed effects to control for unobserved characteristics

In the standard hedonic model[3], price-determining characteristics are explicitly included in the model. In the case of the rental survey, however, the only available characteristics are fine-level region and number of bedrooms. These are not a sufficient measure of the quality of rental dwellings to provide a fully quality-adjusted index when included in the hedonic model. So the hedonic formulation used to benchmark the rental index took advantage of the longitudinal nature of the rental data by fitting dwelling-specific intercepts. That is, a fixed effects model (Allison, 2005) was used to control for all characteristics of the rental dwellings fixed across time.

5.5 Validity of the fixed effects approach

Questions were raised about the appropriateness of using fixed effects in the hedonic model. In particular, does the fixed effects price index implicitly impute a price movement of zero for newly rented dwellings? If so, the benchmarking hedonic index might suffer from downwards bias in a similar way that the matched-sample index might.

To give some context to these concerns, note that the continuously maintained probability sample of the rental survey is unusual in a price measurement context, where it is common for newly-introduced products to be directly replacing old products in a fixed basket. If new rental dwellings in the sample were one-to-one replacements for rentals leaving the sample, then it does seem reasonable to hypothesise that a fixed effects model – which models the difference from the mean rent across time of each dwelling – might be 'zeroing out' any price difference be-

[3] For example as described in section 4 for the estimation of price movements for used-cars

tween old and new dwellings. In other words, implicitly attributing any price difference between the old and new rental to quality change.

To explicitly address the validity of the fixed effects formulation, Krsinich (2011b) extended a result from Aizcorbe, Corrado & Doms (2003) to show that the implicit price movement for a newly rented dwelling is appropriate: it is the movement from the average quality-adjusted[4] rents of continuing[5] dwellings in the previous quarter to the quality-adjusted rent of the new dwelling in the current quarter.
In other words, the price movement implicitly contributed to the index by the new rental dwelling is fully quality-adjusted and as such reflects only price change, with respect to the characteristics that are constant across time at the rental dwelling level..

The properties of this fixed effects price index have since been looked at more generally in the context of scanner and online data by Krsinich (2013, 2014) and de Haan (2013) and this is discussed further in sections 6 and 7.

6 Quality-adjusted price indexes from scanner data

Many products, such as products purchased at supermarkets, have their barcodes scanned at the time of purchase. This retail transaction data – or 'scanner data' – records prices, quantities sold, and associated information for all transactions across the full reference period.

From the 2006 Consumers Price Index Review onwards, aggregated scanner data for supermarket products and for consumer electronics has been used to:

- determine the expenditure weights of some goods in the CPI basket
- determine whether expenditure weight adjustments are required to reflect volume changes since the weight reference period (but before implementation of reviews) and, if so, by how much
- select representative products to survey when price collectors visit retail outlets each month or quarter
- ensure that the mix of brands in the CPI price samples reflect market shares.

[4] With respect to the characteristics of the dwellings that are fixed across time, as these are the characteristics controlled for by the fixed-effects formulation.

[5] Continuing rental dwellings are those that are neither births nor deaths in a given quarter.

Since 2008, Statistics NZ has been actively researching the further potential of using more-detailed scanner data for directly estimating price change for products sold through supermarkets, and for consumer electronics products. The focus has been on determining appropriate methodologies to produce quality-adjusted price indexes.

A detailed research dataset of scanner data for consumer electronics products, from market research company GfK, was used for this purpose.

The potential benefits of using scanner data to measure price change include:

- improved accuracy, due to greater coverage of transactions and availability of real-time quantities
- the ability to use existing administrative-type data sources
- improved treatment of seasonal commodities
- the ability to account for commodity and product substitution between reweights.

However, the behaviour of prices and quantities at the detailed product specification level, as reflected by scanner data, mean that traditional index methods do not work well.

There are two reasons for this:

- high product churn – which makes fixed-base indexes such as the Laspeyres quickly unrepresentative between rebases (See Section 6.1 below), and
- the volatile nature of prices and quantities due to discounting, seasonality, and the life-cycle of products – which may lead to chain-drift in chained indexes (See Section 6.2 below).

6.1 Product churn

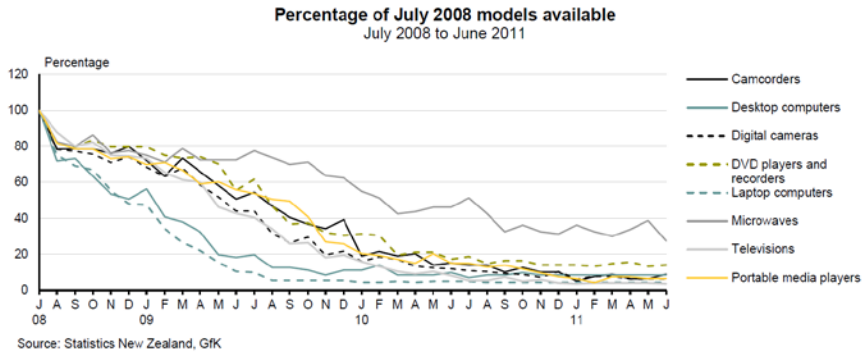
'Churn' refers to the turnover in products[6] being sold in the market. For some product groups, in particular consumer electronics, this turnover can be significant. Figure 3 shows the percentage of products sold in July 2008 that are still on sale each month over the subsequent three years, for the eight consumer electronics goods investigated. For laptop computers, only around 10 percent of the July 2008 products were still being sold a year later.

Note that current CPI practice for consumer electronics is to regularly update products being priced during the time between the three-yearly expenditure weight updates, with manual quality adjustment using a range of methods specific to the products. For example, option costs for computers.

---

[6] That is, products of a given specification, e.g. as reflected by barcodes.

Figure 3. Percentage of July 2008 product specifications still available from July 2008 to June 2011



Given this high degree of product churn, it might seem that one way of maintaining the representativeness of the basket being priced would be to use a chained superlative[7] index. That is, the index is calculated each period in reference to the products available in the current and previous period (rather than the current and base period) and that index number is linked on to the previous index number. This would maximise the number of matched products included in the index calculation and reflect substitution across products by incorporating updated quantities each period. However, this is problematic, as explained in the next section.

6.2 Chain drift caused by price and quantity volatility

High-frequency chained superlative indexes can result in substantial bias - known as 'chain drift' - when applied to scanner data. Ivancic, Diewert, & Fox (2011) show this for Australian supermarket scanner data by a comparison with their rolling year Gini-Elteto-Koves-Szulc (RYGEKS) method[8], which is free of chain-drift by construction.

Chain drift is the bias that occurs when a chained index diverges, or systematically 'drifts' away, from its direct (i.e. unchained) counterpart. A chained index in which the return of prices and quantities to previous levels does not correspond to the index also returning to its previous level, is exhibiting chain drift.

---

[7] A superlative index, such as a Törnqvist index, is one that utilises both current- and reference-period quantity shares symmetrically, which results in substitution between products being accounted for in the index appropriately.

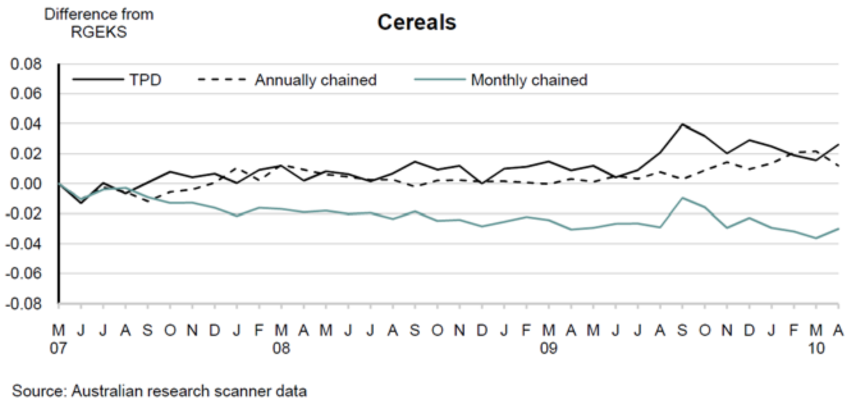
[8] Discussed in more detail in section 6.4

Supermarket products tend to be discounted frequently and, as might be expected, quantities bought increase sharply in response to these discounts as customers stockpile.

In 2010, Statistics NZ and Statistics Netherlands were given access to Australian supermarket scanner research data as peer-reviewers of the Australian Bureau of Statistics' research in this area. Krsinich (2011a) showed that, in general, the monthly chained Törnqvist was biased downwards relative to the RYGEKS for the products considered.

For example, figure 4 shows the results for cereals. Actual price indexes cannot be shown for reasons of confidentiality, so instead the results are presented in terms of difference from the RYGEKS[9]. Note that the time product dummy (TPD) index is the more usual name (in the Prices context) for the fixed effects index[10],

Figure 4. Different methods on Australian scanner data



6.3 Methodologies for scanner data

[9] Referred to in the graph as 'RGEKS' for 'rolling GEKS'.
[10] Called 'time-product dummy' because there are dummy variables included in the model for both time and products, as distinct from the more standard 'time-dummy' (TD) hedonic index which has dummies for time along with explicit inclusion of product characteristics in the model, e.g. as for used cars, discussed in section 4.

Statistics NZ has collaborated with Statistics Netherlands by empirically testing, on New Zealand consumer electronics scanner data, a new benchmark index method called the Imputation Törnqvist Rolling Year GEKS (ITRYGEKS). This method extends the RYGEKS method, which had previously been regarded as the benchmark method for scanner data. De Haan and Krsinich (2014a) show that the RYGEKS, although free of chain-drift, can be biased by not accounting for the implicit price movements associated with new and disappearing products. The ITRYGEKS method incorporates hedonic modelling into the RYGEKS approach in such a way that new and disappearing products are dealt with appropriately.

The ITRYGEKS method is a feasible method for the consumer electronics scanner data, which has extensive information on product characteristics available to incorporate into the hedonic models that the ITRYGEKS method utilises.

However, for supermarket scanner data, the ITRYGEKS method is not feasible because there is insufficient information on product characteristics available in the data. Research is ongoing, but the fixed effects approach first used in the rental benchmarking exercise[11] looks promising (Krsinich, 2013), because the scanner data is longitudinal at the detailed product (i.e. barcode) level, with price-determining characteristics fixed across time by definition.[12]

6.4 The rolling year GEKS

Ivancic, Diewert, and Fox (2011) proposed the rolling year GEKS (RYGEKS) method for producing price indexes from scanner data. The RYGEKS uses all the prices and quantities in the data, and is free of chain drift. It is an extension to the time domain of the Gini, Eltetö and Köves, and Szulc (GEKS) method used for calculating multilateral spatial price indexes such as purchasing power parities.[13]

Within some window of time (generally one year plus a period, i.e. five quarters or 13 months[14]) the GEKS index between periods r and s is the geometric mean of all the superlative[15] bilateral indexes between:

[11] See section 5.

[12] As any change in a product's characteristics will mean the barcode changes.

[13] Purchasing power parities compare prices in different countries at a given point in time

[14] The choice of a year plus a period means that the window will be long enough to counter periods of seasonal unavailability. This is the generally accepted default window length, but more research is required to formulate window-length optimality and the concepts defining it.

[15] Any superlative index could be used (e.g. Fisher, Tornqvist, Walsh) but Ivancic et al (2011) use the Fisher index.

- time r and all the other periods in the window, and
- time s and all the other periods in the window.

Formulating the monthly RYGEKS with a 13-month rolling window is as follows:

For the first window, ie t=0 to 12, the RYGEKS index is equal to the GEKS index:

$$P_{RYGEKS}^{OT} = P_{GEKS}^{OT} = \prod_{t=0}^{T} \left[P^{0t} \times P^{tT} \right]^{1/(T+1)} \qquad (3)$$

Where P^{rs} is any superlative index (e.g. the Törnqvist) between periods r and s.

From t=13 onwards, the RYGEKS links on the most-recent movement from the GEKS calculated on the next window (ie from t=1 to 13, then from t= 2 to 14, and so on) as follows:

$$P_{RYGEKS}^{0,13} = P_{GEKS}^{0,12} \prod_{t=1}^{13} \left[P^{12,t} \times P^{t,13} \right]^{1/13} = \prod_{t=0}^{12} \left[P^{0,t} \times P^{t,12} \right]^{1/13} \prod_{t=1}^{13} \left[P^{12,t} \times P^{t,13} \right]^{1/13}$$

and then

$$P_{RYGEKS}^{0,14} = P_{RYGEKS}^{0,13} \prod_{t=2}^{14} \left[P^{13,t} \times P^{t,14} \right]^{1/13} \qquad (4)$$

and so on.

A limitation of the RYGEKS method, however, is that it is based on the price movements between matched[16] products only. Any implicit price change associated with new or disappearing products is not reflected. So, for example, if the initial price of the latest model of a mobile phone is high relative to its features when it is first introduced then this implicit price increase is not reflected in the RYGEKS.

6.5 The imputation Törnqvist rolling year GEKS

[16] That is, existing in both the two periods relating to the individual bilateral indexes that feed into the RYGEKS

An extension to the RYGEKS method was proposed by Jan de Haan, of Statistics Netherlands (de Haan and Krsinich, 2014a). The imputation Törnqvist rolling year GEKS (ITRYGEKS) uses hedonic models to impute for new and disappearing products. Unlike the RYGEKS, which is based on the geometric means of superlative bilateral indexes, the ITRYGEKS is based on geometric means of hedonic bilateral indexes.

The formulation is as above for RYGEKS, with the difference that the P^{ij} are bilateral time-dummy hedonic indexes.

De Haan and Krsinich (2014a) show that, with appropriate weights[17], the ITRYGEKS is algebraically equivalent to a Törnqvist index for the matched subset of products[18]. For new and disappearing products, the ITRYGEKS is algebraically equivalent to applying a Törnqvist formula to prices predicted from time-dummy hedonic models for the period in which there is no price available.

That is, the ITRYGEKS from period 0 to t can be expressed as follows:

$$P_{ITRYGEKS}^{Ot} = \prod_{i \in M^{0t}} \left(\frac{p_i^t}{p_i^0} \right)^{\frac{s_i^0 + s_i^t}{2}} \prod_{i \in D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0} \right)^{\frac{s_i^0}{2}} \prod_{i \in N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0} \right)^{\frac{s_i^t}{2}} \tag{5}$$

Where

M^{0t} is the set of matched products with respect to periods 0 and t
D^{0t} is the set of 'disappearing' products with respect to periods 0 and t – that is, the subset of products that exist in period 0 but not in period t
N^{0t} is the set of 'new' products with respect to periods 0 and t – that is, the subset of products that exist in period t but not in period 0
p_i^t is the actual price for product i in period t
$\hat{p}_i^t$ is the predicted price for product i in period t from the time dummy hedonic model based on the pooled bilateral data of periods 0 and t
s_i^t is the expenditure share for item i in period t

[17] The mean expenditure shares are the weights for the matched items, and half of the expenditure shares are the weights used for the unmatched products

[18] By ‘products’ here, we mean products with a particular specification, that is, at the level of the barcode.

And this expression can be generalised to the ITRYGEKS between any two periods r and s.

The ITRYGEKS method was empirically tested using scanner data for eight New Zealand consumer electronics products (camcorders, desktop computers, digital cameras, DVD players and recorders, laptop computers, microwaves[19], televisions and portable media players) resulting in the following conclusions:

- The monthly chained Törnqvist is not a viable method for consumer electronics products, as it exhibits downwards chain drift for most products examined.
- RYGEKS shows evidence of bias due to not accounting for the price movements of new and disappearing products, particularly for computers.
- The easier-to-implement[20] rolling year time-dummy (RYTD) hedonic index based on multilateral pooled[21] data gives very similar results to the ITRYGEKS.
- In some cases, such as supermarket data, few or no characteristics are likely to be available for explicitly incorporating into time dummy hedonic models. Therefore neither the ITRYGEKS or RYTD methods are feasible. The fixed effects method (called the rolling year time-product dummy (RYTPD) method in the paper) sits closer to the benchmark ITRYGEKS than does the RYGEKS, and therefore should be further investigated.

6.6 Fixed effects

As noted above, ITRYGEKS is a feasible method for production only when there are sufficient characteristics in the data to estimate the underlying bilateral time dummy hedonic indexes. The research data obtained for consumer electronics has extensive characteristics data so can support the estimation of ITRYGEKS indexes.

[19] Microwaves are not strictly a consumer electronics product but, as a product with relatively little technological change, they offer a useful comparison.

[20] But less able to be expressed in terms of traditional price index formulae

[21] That is, the hedonic models are based on the data for the full estimation window of, for example, 13 months.

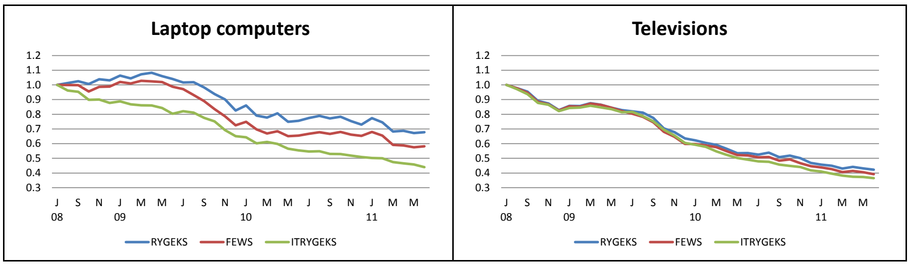
However, for supermarket data, the only explicit characteristics likely to be consistently available in the data across all products are weight, volume, or size. This means ITRYGEKS will not be feasible for supermarket products.

Methods which don't require explicit information on characteristics are chained superlative indexes such as the Tornqvist (which has been shown to suffer from chain drift as discussed above), the RYGEKS and the fixed effects index (also known as the time-product dummy, or TPD method).

De Haan and Krsinich (2014) showed empirically that the fixed effects index tends to sit closer to the ITRYGEKS benchmark than the RYGEKS approach does. However they noted that, by construction, price movements of new items are systematically omitted due to the splicing, or 'linking on', of only the most recent period's movement. Fixed effects regression requires two price observations before a product is non-trivially included in the estimation. Krsinich (2013) proposed and empirically tested a modified version of the splicing procedure – where the movement of the entire estimation window is incorporated rather than just the most recent period's movement. This approach ensures that the price movements of new products will be reflected in the index, in the period of entry, with one period's lag. This was shown to be consistently closer to the ITRYGEKS than either the RYGEKS or the fixed effects index which splices on just the most recent period's movement – i.e. the RYTPD empirically tested in de Haan and Krsinich (2014).

Figure 4 shows this fixed effects window-splice (FEWS) index compared to the RYGEKS and the ITRYGEKS, for laptop computers and televisions.

Figure 4. Methods requiring no information on characteristics, compared with the benchmark ITRYGEKS



The work on the rentals fixed effects benchmark, described above in section 5, showed that the implicit price movement being estimated for new rental dwellings is appropriate, and the results on the consumer electronics scanner data show that the method performs better than the RYGEKS with respect to the ITRYGEKS benchmark.

Intuitively, though, it is difficult to see why this should be the case. How can the fixed effects index work more effectively than the RYGEKS, if there is no extra information being incorporated? As with the RYGEKS, only prices, quantities and detailed product identifiers are utilised.

In fact there *is* extra information being utilised by the fixed effects method – the longitudinal aspect of the data. Unlike the RYGEKS, the full longitudinal record for each product (once there are at least two price observations) is modelled in the fixed effects index. In the RYGEKS, only 2-period, or bilateral, movements of matched products are incorporated into the estimation.

Krsinich (2014) extends a result from de Haan (2013) to show that the fixed effects index is equivalent to a fully-interacted[22] time-dummy hedonic index, where all the characteristics are incorporated into the model as categorical variables.

In other words, the fixed effects index may be a *better* quality-adjusted index than the time-dummy hedonic index (including only main effects) which explicitly incorporates information on characteristics. In part, because all time-invariant characteristics[23] are (implicitly) controlled for, not only the observed characteristics. Also, because all the interactions between these characteristics are (implicitly) incorporated into the model.

This is a new idea which is still being debated, but it holds significant potential for very simple and effective quality-adjustment of price indexes from 'big data' such as scanner or online data where information on characteristics is limited or non-existent, but where product specifications are at a level of detail that ensures constant characteristics over time in the longitudinal data.

Very recent research by de Haan and Krsinich (2014b) utilises the New Zealand consumer electronics scanner data to show that the time-dummy hedonic index (i.e. where characteristics are explicitly included in the model) corresponds very closely to a 'quality-adjusted unit value' (QAUV) index, defined according to established econometric principles (de Haan, 2004). If the fixed effects index is equivalent to a fully interacted time-dummy hedonic index then it will follow that it can similarly be related to the concept of a quality adjusted unit value index.

---

[22] That is, a time-dummy hedonic index where the hedonic model includes all the interactions between characteristics, as well as the main effects.

[23] Note that all characteristics at the barcode level are time-invariantby definition, because a change in any of them will result in a changed barcode.

7 Quality adjusted price indexes from online data

Statistics New Zealand has recently started to consider the potential of web-scraped online retail data for constructing price indexes. We have done some initial exploratory analysis on 15 months of daily online prices for a major New Zealand consumer electronics retailer. This data was shared with us by the Billion Prices Project of MIT, an academic initiative that uses prices collected from hundreds of online retailers around the world on a daily basis to conduct economic research. See Cavallo (2012, b) for a description of how inflation measures based on the online data were first used to find evidence of bias in the official Argentine CPI measure.

Online data has product identifiers at the model level – similar to scanner data – but will tend to have little, if any, data on the characteristics of products. It also lacks any information on the quantities sold.

The question then is how to produce quality-adjusted indexes from this data. Daily inflation measures marketed by PriceStats[24] for over 20 countries use a daily chained Jevons index. This is a matched-model approach to quality adjustment where the price movement is calculated based on the products that are available in the previous and current day. The absence of quantities in the data means that there is likely to be less potential for chain drift[25] but, as with any matched-model approach, the implicit price movements associated with new products appearing and old products disappearing won't be reflected in the index. Cavallo (2012a) examined this issue by using overlapping prices to adjust for quality change. Lacking data on quantities is also a problem because the indexes will not fully reflect different expenditure shares across products, and the changes in these[26].

Statistics New Zealand is investigating the potential of the fixed effects approach in the context of this online data, and initial results are promising.

---

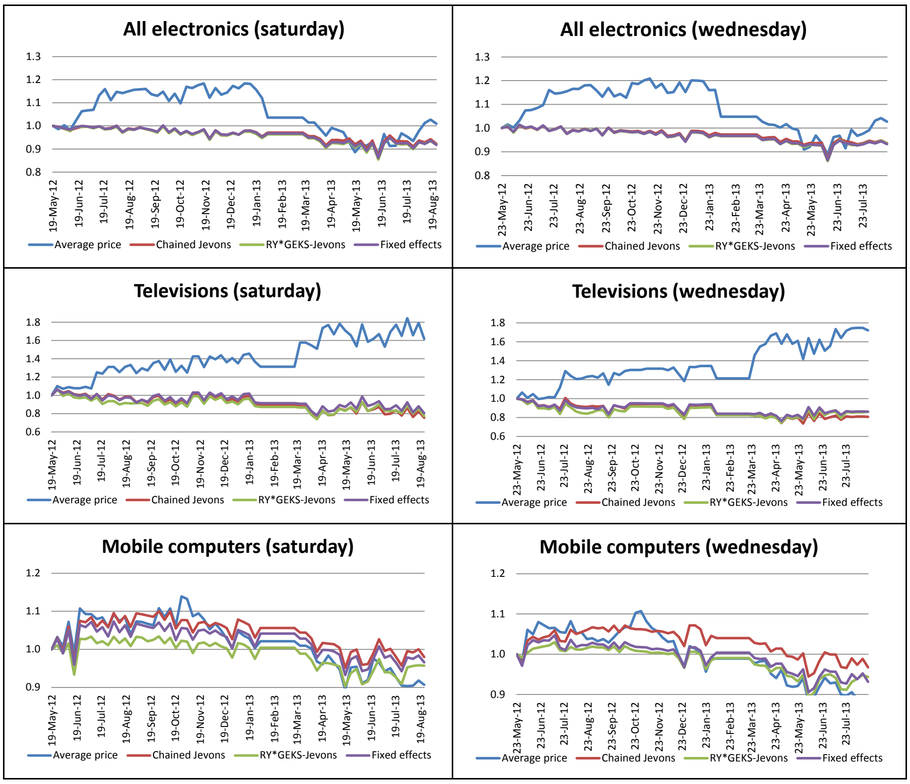
[24] PriceStats is a US company which is related to, but distinct from, the Billion Prices Project academic initiative of MIT. It is a US-based company that seeks to bring to market innovation in the inflation measurement field. See http://www.pricestats.com/

[25] As it appears to be the slightly assymetrical behavior of the price and quantity spikes that causes chain drift, though the author is unsure whether this has been formally established in the literature.

[26] Note that at the higher levels of aggregation in the CPI structure, the PriceStats inflation measures utilize expenditure weightings from the relevant countries' CPIs.

Figure 5 shows indexes with no quality adjustment (the 'average price' graphs) compared to RYGEKS[27], weekly chained Jevons and fixed effects indexes. Note that the data collection was interrupted during February and March 2013, so the prices were carried forward over this period (hence the indexes are flat for these 2 months). Because of the difficulty of estimating RYGEKS for very high frequency data due to the very large number of bilateral index calculations required, we took day-per-week samples of the daily data, and computed the indexes for each of these. This is interesting in itself, as it shows the effect of sampling in the time-dimension which, in this case, can have an impact on estimates of short term movements, though not the longer-term trend.

Figure 5. Different methods on online data



[27] This RY*GEKS is a modification of the usual approach that would be used in production, where a year-long GEKS index would be recalculated each period (e.g. day or week, depending on the frequency of the index). Instead, two 53-week GEKS are linked together in the middle of the 3 year period.

While for televisions the three quality-adjusted index methods are very similar, for mobile computers they can differ quite significantly, reflecting the potential for chain drift in the chained Jevons and bias due to not accounting for the price change of new goods appearing and old goods disappearing, in both the RYGEKS and the chained Jevons.

At the aggregate consumer electronics level, the methods give very similar results, suggesting a certain amount of 'cancelling out' of biases across products.

This initial analysis indicates that the fixed effects method may be an appropriate one for producing quality-adjusted indexes from online data[28]. Statistics New Zealand is in the beginning stages of a collaborative research agreement with PriceStats to share their daily online data for a range of New Zealand retailers who sell their products online.

8 Conclusion

Ensuring that the New Zealand CPI reflects 'like-for-like' price change is important. Products such as consumer electronics are rapidly evolving and have correspondingly shorter life-cycles in the market. This makes the traditional approach of pricing a representative fixed basket challenging.

Hedonic regression has been used in production in the NZ CPI for over 10 years, to measure quality-adjusted price movements of used cars. The hedonic approach was also recently used to benchmark the performance of the current matched-sample approach to measuring housing rentals. This was the first example of Statistics New Zealand exploiting the longitudinal nature of the data to control for unobserved price-determining characteristics using fixed effects.

New data sources, in particular scanner data and web-scraped online data, have significant potential benefits, but they require new methodologies to ensure that the price indexes derived from them are appropriately quality-adjusted. This has been an active area of research for Statistics New Zealand over the last five years. Collaborative research with Statistics Netherlands on a method extending the RYGEKS method to incorporate hedonic indexes has been well-received in the Prices research community, and the imputation Tornqvist RYGEKS (ITRYGEKS) may now be considered a benchmark method for quality-adjusting price indexes from scanner data. Statistics New Zealand intends to incorporate scanner data into production for the NZ CPI in the September quarter of 2014, for consumer electronics products, using the ITRYGEKS method.

[28] The issue remains that online data lacks information on quantities sold, though existing CPI weighting information could be used to aggregate the product-level indexes at higher levels of the index hierarchy.

Supermarket scanner data and web-scraped online data, however, do not contain sufficient information on characteristics to use the ITRYGEKS method. This has led to a revisiting of the fixed effects index, along with a modification of the standard splicing procedure to ensure the index is non-revisable and also reflects the price movements of new products appropriately. At time of writing, this looks very promising as an approach to producing non-revisable fully quality-adjusted price indexes where there is longitudinal price and quantity information at a detailed product specification level. Research in this area is ongoing.

References

Allison, P D (2005). Fixed effects regression methods for longitudinal data using SAS. Cary, NC: SAS Institute Inc.

Aizcorbe, A, Corrado, C, & Doms, M (2003). When do matched-model and hedonic techniques yield similar price measures? Working Paper no. 2003-14, Federal Reserve Bank of San Francisco.

Boskin, M, Dulberger, E, Gordon, R, Griliches. Z, & Jorgenson, D (1996, December). Toward a more accurate measurement of the cost of living: Final report to the Senate Finance Committee. Washington: Advisory commission to study the consumer price index.

Cavallo, A (2012a). Overlapping quality adjustment using online data, Paper presented at the 2012 Economic Measurement Group, Sydney, Australia.

Cavallo, A (2012b). Online versus official price indexes: Measuring Argentina's inflation. Journal of Monetary Economics. December 2012.

de Haan, J (2004). Estimating quality-adjusted unit value indexes: Evidence from scanner data. Paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, 30 June – 3 July 2004, Vancouver, Canada.

de Haan, J (2013). Online data, fixed effects and the construction of high-frequency indexes, Paper presented at the 2013 Economic Measurement Group, Sydney, Australia. November 2013.

de Haan, J. and F Krsinich (2014a) Scanner data and the treatment of quality change in non-revisable price indexes, Journal of Business & Economic Statistics (forthcoming) de Haan, J. and F Krsinich (2014b) Time dummy hedonic and quality-adjusted unit value indexes: do they really differ?, to be presented at the inaugural conference of the Society of Economic Measurement, Chicago, August 2014

International Labour Organisation (2003). Resolution concerning consumer price indices. Available from www.ilo.org.

Ivancic, L, Diewert, WE, & Fox, KJ (2011). Scanner data, time aggregation and the construction of price indexes, Journal of Econometrics, 161(1), 24–35.

Krsinich, F (2009) Use of hedonics at Statistics New Zealand. Paper presented at the 2009 Economic Measurement Group. Sydney, Australia. December 2009.

Krsinich, F (2011a, May). Price indexes from scanner data: A comparison of different methods, Paper presented at the 12th meeting of the Ottawa Group, Wellington, New Zealand.

Krsinich, F (2011b, May). Measuring the price movements of used cars and residential rents in the New Zealand consumers price index, Paper presented at the 12th meeting of the Ottawa Group, Wellington, New Zealand.

Krsinich, F (2011c, May). An assessment of New Zealand's unofficial SPAR and stratified house price methods, Room document tabled at the 12th meeting of the Ottawa Group, Wellington, New Zealand.

Krsinich, F (2013) Using the rolling year time-product dummy method for quality adjustment in the case of unobserved characteristics, Paper presented at the 2013 Ottawa Group, Copenhagen, Denmark. May 2013.

Krsinich, F (2014) Fixed effects with a window splice: quality-adjusted price indexes with no characteristic information. Paper to be presented at the 2014 UNECE/ILO meeting of price index experts, Geneva, Switzerland. May 2014.

Melser, D (2011, May). Constructing cost of living indexes using scanner data, unpublished draft, September 2011, an updated version of Constructing high frequency indexes using scanner data, Paper presented at the 12th meeting of the Ottawa Group, Wellington, New Zealand.

Smith, M (2008). Measuring rents in the CPI. Unpublished document, Reserve Bank of New Zealand.

New Zealand Government (1912) Cost of Living in New Zealand (report and evidence of the Royal Commission on). Department of Labour. Appendices to the Journal of the House of Representatives (AJHR) H–18

New Zealand Government (1948). Report of index committee. AJHR H–48.

New Zealand Government (1978). Report of the consumers price index revision advisory committee. Oct 1978. AJHR G. 28A.

New Zealand Government (1985). Report of the consumers price index revision advisory committee. May 1985. AJHR G. 28A.

New Zealand Government (1991). Report of the consumers price index revision advisory committee. Nov 1991. AJHR G.28A.

Smith, M (2008). Measuring rents in the CPI. Unpublished document, Reserve Bank of New Zealand.

Statistics NZ (2011a). Bigger, safer, better: tracking retail and quality adjusted new car prices in the CPI. http://www.stats.govt.nz/browse_for_stats/economic_indicators/prices_indexes/new-car-prices.aspx

Statistics NZ (2011b). Updating the CPI used cars index hedonic model. http://www.stats.govt.nz/browse_for_stats/economic_indicators/prices_indexes/used-cars-hedonic-model.aspx

Statistics NZ (2012). Accounting for quality change in the CPI. http://www.stats.govt.nz/tools_and_services/newsletters/price-index-news/jul-12-article-cpi-quality-change.aspx

Statistics NZ (2013). Retail transaction data. Chapter 5 of the discussion paper for the CPI advisory committee 2013. http://www.stats.govt.nz/browse_for_stats/economic_indicators/CPI_inflation/2013-cpi-review-advisory-committee.aspx#discussion

Acknowledgments

The author gratefully acknowledges the assistance and encouragement of Chris Pike, Jan de Haan, Alistair Gray, Soon Song, Sharleen Forbes and Richard Arnold.

Alan Bentley was the lead author of an earlier article (Statistics New Zealand, 2012) which formed the basis of section 3.

Also we are very grateful to Alberto Cavallo of MIT for sharing Billion Prices Project online data with us for exploratory research.