

The Digital Economy, New Products and Consumer Welfare

W. Erwin Diewert, Kevin J. Fox and Paul Schreyer

ESCoE Discussion Paper 2018-16

November 2018

ISSN 2515-4664



About the Economic Statistics Centre of Excellence (ESCoE)

The Economic Statistics Centre of Excellence provides research that addresses the challenges of measuring the modern economy, as recommended by Professor Sir Charles Bean in his Independent Review of UK Economics Statistics. ESCoE is an independent research centre sponsored by the Office for National Statistics (ONS). Key areas of investigation include: National Accounts and Beyond GDP, Productivity and the Modern economy, Regional and Labour Market statistics.

ESCoE is made up of a consortium of leading institutions led by the National Institute of Economic and Social Research (NIESR) with King's College London, innovation foundation Nesta, University of Cambridge, Warwick Business School (University of Warwick) and Strathclyde Business School.

ESCoE Discussion Papers describe research in progress by the author(s) and are published to elicit comments and to further debate. Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the ESCoE, its partner institutions or the ONS.

For more information on ESCoE see www.escoe.ac.uk.

Contact Details

Economic Statistics Centre of Excellence
National Institute of Economic and Social Research
2 Dean Trench St
London SW1P 3HE
United Kingdom

T: +44 (0)20 7222 7665
E: escoeinfo@niesr.ac.uk



The Digital Economy, New Products and Consumer Welfare

W. Erwin Diewert^{1,2}, Kevin J. Fox² and Paul Schreyer³

¹ University of British Columbia

² UNSW Sydney

³ OECD

Abstract

Benefits of the Digital Economy are evident in everyday life, but there are significant concerns that these benefits may not be appropriately reflected in official statistics. The measurement of the net benefits of new and disappearing products depends on what type of price index the statistical agency is using to deflate final demand aggregates. We examine this measurement problem when the agency uses any standard price index formula for its deflation of a value aggregate, such as GDP. We also apply the methodology to the problem of measuring the effects of product substitutions for disappearing items. Our exact expressions for biases inherent in different approaches provide a theoretical basis and framework for the emerging empirical literature on new goods and services, and for assessing the quality adjustment methodologies used in practice.

Keywords: Maximum overlap indexes, Hicksian reservation prices, quality adjustment, replacement sampling, index number bias.

JEL classification: C43, D11, D60, E01, E31, O31, O47

Contact Details

Kevin J. Fox
School of Economics & CAER
UNSW Sydney
NSW 2052
Australia

Email: erwin.diewert@ubc.ca, K.Fox@unsw.edu.au, paul.schreyer@oecd.org

This ESCoE paper was first published in November 2018.

© W. Erwin Diewert, Kevin J. Fox and Paul Schreyer

The Digital Economy, New Products and Consumer Welfare

W. Erwin Diewert

University of British Columbia and UNSW Sydney

Kevin J. Fox*

UNSW Sydney

Paul Schreyer

OECD

November 2018

Abstract

Benefits of the Digital Economy are evident in everyday life, but there are significant concerns that these benefits may not be appropriately reflected in official statistics. The measurement of the net benefits of new and disappearing products depends on what type of price index the statistical agency is using to deflate final demand aggregates. We examine this measurement problem when the agency uses any standard price index formula for its deflation of a value aggregate, such as GDP. We also apply the methodology to the problem of measuring the effects of product substitutions for disappearing items. Our exact expressions for biases inherent in different approaches provide a theoretical basis and framework for the emerging empirical literature on new goods and services, and for assessing the quality adjustment methodologies used in practice.

Keywords: Maximum overlap indexes, Hicksian reservation prices, quality adjustment, replacement sampling, index number bias.

JEL Classification Numbers: C43, D11, D60, E01, E31, O31, O47

***Corresponding Author:** Kevin J. Fox, School of Economics & CAER, UNSW Sydney, NSW 2052, Australia. E-mail: K.Fox@unsw.edu.au, Tel: +61-2-9385-3320. Other contacts: erwin.diewert@ubc.ca; paul.schreyer@oecd.org. We thank two anonymous reviewers, from ESCoE and the UK Office for National Statistics, Dennis Fixler, Jan de Haan, Michael Webster, and participants at the ESCoE Conference on Economic Measurement (Bank of England, May 2018) and the WEAI Conference (Vancouver, June 2018), for helpful comments. The first author gratefully acknowledges the financial support of the SSHRC of Canada, and the first two authors gratefully acknowledge the financial support of the Australian Research Council (DP150100830). The views expressed in this paper reflect those of the authors and not necessarily those of the OECD or its Member countries.

1 Introduction

The debate regarding the benefits of the Digital Economy regularly features the suggestion that mismeasurement by national statistical offices is playing a major role in obscuring productivity and economic growth, price declines and welfare gains.¹ If measurement is lacking, through methodological challenges, statistical agency budgets or data availability, then we are severely hampered in our ability to understand the impact of new technologies, goods and services. In this paper, we build on an earlier attempt by Diewert and Fox (2017a) and develop new frameworks for measuring welfare change and real consumption growth in the presence of new and disappearing goods (and services); such goods are frequently synonymous with the Digital Economy.

New, often very specialized digital goods are now part of daily consumption for many, accompanied by the disappearance of previously consumed commodities. In addition, digital product design and sales platforms allow rapid product turnover. For example, using Adobe Analytics data on online transactions from January 2014 to September 2017, Goolsbee and Klenow (2018; 490) found that “roughly half of the sales volume online is for products that did not exist in the previous year. Even without apparel, the figure is 44%. The products that disappear, meanwhile, had about 24% of total sales before they left the market (22% excluding apparel).”

In this paper we provide a framework for quantifying the welfare benefits and costs of new and disappearing products.² The basic idea is the following. Statistical agencies typically use a “matched model” approach when they construct price indexes, and these are used to deflate a final demand value aggregate; i.e., when constructing a particular price index that compares the prices of a group of products over two periods, the scope of the index is usually restricted to the set of commodities or products that are present in both periods. The resulting index is called a *maximum overlap index*.³ However, if one is using the economic approach to index number theory that was originally developed by Konüs (1924), then *reservation prices* for the missing products should be matched up with the zero quantities for the missing products in each period; the reservation price for a

¹See, for example, Goodridge, Haskel and Wallis (2015), Bean (2016), Ahmad and Schreyer (2016), Byrne, Fernald and Reinsdorf (2016), Feldstein (2017), Reinsdorf and Schreyer (2017), Syverson (2017), Groshen, Moyer, Aizcorbe, Bradley and Friedman (2017), Ahmad, Ribarsky and Reinsdorf (2017), and Hulten and Nakamura (2017).

²Diewert and Fox (2017a) did not consider the case of disappearing goods, but considered the case of free goods. While free goods often have an implicit price, this price is usually unobserved. In this case, a price of zero is applied by national statistical offices, resulting in the positive quantities of these goods having no measured value. Hence their benefits to consumers go unmeasured, and they do not appear in nominal and real output. See Brynjolfsson, Eggers and Gannamaneni (2018), Brynjolfsson, Diewert, Eggers, Fox and Gannamaneni (2018) and Brynjolfsson and Oh (2012) on measuring the value of free digital services, and Nakamura, Samuels and Soloveichik (2016) for examples of how to think about the valuation of free media.

³This type of index dates back to Marshall (1887). Keynes (1930; 94) called it the highest common factor method while Triplett (2004; 18) called it the overlapping link method. See Diewert (1993a; 52-56) for additional material on the early history of the new goods problem.

missing product is the price which would induce a utility maximizing potential purchaser of the product to demand zero units of it. Normal index number theory can then be applied to the resulting augmented data set for the two periods under consideration.⁴

This reservation price approach for the treatment of new goods is due to Hicks (1940; 114).⁵ Hofsten (1952; 95-97) extended his approach to cover the case of disappearing goods. If reservation prices are estimated, elicited from surveys,⁶ or guessed, then the “true” price index can be calculated and compared to its maximum overlap counterpart. Feenstra (1994) uses our suggested general methodological approach in the context of purchasers who have CES preferences, and he uses the Sato (1976)-Vartia (1976) maximum overlap index number formula which is exact for CES preferences. An advantage of his methodology is that the Hicksian reservation prices in the CES context are equal to $+\infty$ and thus there is no need to estimate these reservation prices. However, $+\infty$ is a high reservation price, implying a large price decline when the product becomes available. Hence it is likely that his approach overestimates the benefits of new products; this is the approach used by e.g. Goolsbee and Klenow (2018) in estimating the impact of new products on inflation using Adobe Analytics data. Moreover, the CES functional form is not fully flexible, in contrast to the preferences that are exact for the Törnqvist and Fisher indexes; see Diewert (1976).⁷

Using our approach, an estimate of the bias in the maximum overlap deflator can be formed. This translates into a corresponding bias in the growth of the real output aggregate, which from index number theory can be interpreted as a measure of welfare change.⁸

We will evaluate this bias in the context of a statistical agency that uses maximum overlap Törnqvist indexes in section 3. The context we consider is one in which transaction level data are available so that indexes can be calculated from the elementary level. In a similar manner, we will evaluate the bias in the Laspeyres, Paasche and Fisher maximum overlap indexes in section 4. Section 2 develops some general relationships in the expenditure shares of a true index relative to its maximum overlap counterpart. These relationships will be used in sections 3 and 4.

Finally, section 5 applies the algebra developed in sections 3 and 4 to the problem of measuring the effects of product substitutions for disappearing items. Specifically, we iden-

⁴See Diewert (1976) for practical applications of the economic approach to index number theory.

⁵Rothbarth (1941) called these virtual prices. Hicks did not give a name to his pricing concept.

⁶See Brynjolfsson, Eggers, and Gannamaneni (2017) on the use of online choice experiments to elicit valuations of goods and services.

⁷Hausman (1996) (1999) (2003) and Hausman and Leonard (2002) develop an expenditure function approach that uses a flexible functional form to estimate reservation prices. Diewert and Feenstra (2018) evaluate proposes approaches to estimation of reservation prices and suggest some alternatives.

⁸The correspondence between welfare change and quantity indexes is well established. Konüs and Byushgens (1926) showed that for a linearly quadratic utility function, the ratio of utility between two periods is equal to a Fisher quantity index. For more general results, see Allen (1949) and Diewert (1976)(1993b).

tify biases arising from replacement sampling with quality change.⁹ This seemingly basic issue of index construction is becoming increasingly important, as much of the discussion about capturing the impact of product (and process) innovation in the digital economy has focussed on more appropriately adjusting price indexes for quality change.¹⁰

2 The Relationships between True Shares and Maximum Overlap Shares

Consider two periods, 0 and 1. There are three groups of commodities. Group 1 products are present in both periods with positive prices and quantities for all N products in this group. Denote the period t price and quantity vectors for this group of products as $p_1^t \equiv [p_{11}^t, \dots, p_{1N}^t] \gg 0_N$ and $q_1^t \equiv [q_{11}^t, \dots, q_{1N}^t] > 0_N$ for $t = 0, 1$.

Group 2 products are the *new* goods and services that are not available in period 0 but are available in period 1. Denote the period 0 price and quantity vectors for this group of K products as $p_2^{0*} \equiv [p_{21}^{0*}, \dots, p_{2K}^{0*}] \gg 0_K$ and $q_2^0 \equiv [q_{11}^0, \dots, q_{1K}^0] = 0_K$. The prices in the vector p_2^{0*} are the positive reservation prices that make the demand for these products in period 0 equal to zero. These reservation prices have to be estimated somehow. The period 1 price and quantity vectors for these K products are $p_2^1 \equiv [p_{21}^1, \dots, p_{2K}^1] \gg 0_K$ and $q_2^1 \equiv [q_{21}^1, \dots, q_{2K}^1] > 0_K$ and these vectors are observable.

Group 3 products are the *disappearing* goods and services that were available in period 0 but are not available in period 1.¹¹ Denote the period 0 price and quantity vectors for this group of M products as $p_3^0 \equiv [p_{31}^0, \dots, p_{3M}^0] \gg 0_M$ and $q_3^0 \equiv [q_{31}^0, \dots, q_{3M}^0] > 0_M$. The period 1 price and quantity vectors for these M products are $p_3^{1*} \equiv [p_{31}^{1*}, \dots, p_{3M}^{1*}] \gg 0_M$ and $q_3^1 \equiv [q_{31}^1, \dots, q_{3M}^1] = 0_M$. The prices in the vector p_3^{1*} are the positive reservation prices that make the demand for these products in period 1 equal to zero. Again, these reservation prices have to be estimated somehow.

Define the *true expenditure shares* for product n in Group 1 for periods 0 and 1, s_{1n}^0

⁹The problem of not appropriately adjusting for quality change was highlighted by the U.S. Boskin Commission (1996). See also Bills (2009).

¹⁰See for example Bureau of Labor Statistics (2018), Groshen *et al.* (2017), Byrne, Oliner and Sichel (2017), and Byrne and Corrado (2016).

¹¹Diewert and Fox (2017a) did not consider the case of disappearing goods. Another difference between Diewert and Fox (2017a) and the current paper is that here we use a traditional index number approach, whereas the previous paper used an indicator approach, i.e. a difference framework. The results of Diewert and Fox (2017a) are incorporated into Brynjolfsson, Diewert, Eggers, Fox and Gannamaneni (2018).

and s_{1n}^1 , in the usual way:¹²

$$\begin{aligned} s_{1n}^0 &\equiv p_{1n}^0 q_{1n}^0 / [p_1^0 \cdot q_1^0 + p_2^{0*} \cdot q_2^0 + p_3^0 \cdot q_3^0]; \quad n = 1, \dots, N; \\ &= p_{1n}^0 q_{1n}^0 / [p_1^0 \cdot q_1^0 + p_3^0 \cdot q_3^0] \quad \text{since } q_2^0 = 0_N; \end{aligned} \quad (1)$$

$$\begin{aligned} s_{1n}^1 &\equiv p_{1n}^1 q_{1n}^1 / [p_1^1 \cdot q_1^0 + p_2^1 \cdot q_2^1 + p_3^{1*} \cdot q_3^1]; \quad n = 1, \dots, N; \\ &= p_{1n}^1 q_{1n}^1 / [p_1^1 \cdot q_1^1 + p_2^1 \cdot q_2^1] \quad \text{since } q_3^1 = 0_N. \end{aligned} \quad (2)$$

Note that these shares can be calculated using observable data; i.e., these shares do not depend on the imputed prices p_2^{0*} and p_3^{1*} .

Define the *true expenditure shares* for product k in Group 2 for periods 0 and 1, s_{2k}^0 and s_{2k}^1 , as follows:

$$\begin{aligned} s_{2k}^0 &\equiv p_{2k}^0 q_{2k}^0 / [p_1^0 \cdot q_1^0 + p_2^{0*} \cdot q_2^0 + p_3^0 \cdot q_3^0]; \quad k = 1, \dots, K; \\ &= p_{2k}^0 q_{2k}^0 / [p_1^0 \cdot q_1^0 + p_3^0 \cdot q_3^0] \quad \text{since } q_2^0 = 0_K; \\ &= 0; \quad \text{since } q_{2k}^0 = 0; \end{aligned} \quad (3)$$

$$\begin{aligned} s_{2k}^1 &\equiv p_{2k}^1 q_{2k}^1 / [p_1^1 \cdot q_1^0 + p_2^1 \cdot q_2^1 + p_3^{1*} \cdot q_3^1]; \quad k = 1, \dots, K; \\ &= p_{2k}^1 q_{2k}^1 / [p_1^1 \cdot q_1^1 + p_2^1 \cdot q_2^1] \quad \text{since } q_3^1 = 0_K. \end{aligned} \quad (4)$$

Note that these shares can also be calculated using observable data.

Define the *true expenditure shares* for product m in Group 3 for periods 0 and 1, s_{3m}^0 and s_{3m}^1 , as follows:

$$\begin{aligned} s_{3m}^0 &\equiv p_{3m}^0 q_{3m}^0 / [p_1^0 \cdot q_1^0 + p_2^{0*} \cdot q_2^0 + p_3^0 \cdot q_3^0]; \quad m = 1, \dots, M; \\ &= p_{3m}^0 q_{3m}^0 / [p_1^0 \cdot q_1^0 + p_3^0 \cdot q_3^0] \quad \text{since } q_2^0 = 0_M; \end{aligned} \quad (5)$$

$$\begin{aligned} s_{3m}^1 &\equiv p_{3m}^1 q_{3m}^1 / [p_1^1 \cdot q_1^0 + p_2^1 \cdot q_2^1 + p_3^{1*} \cdot q_3^1]; \quad m = 1, \dots, M; \\ &= p_{3m}^1 q_{3m}^1 / [p_1^1 \cdot q_1^1 + p_2^1 \cdot q_2^1] \quad \text{since } q_3^1 = 0_M; \\ &= 0 \quad \text{since } q_{3m}^1 = 0. \end{aligned} \quad (6)$$

Note that these shares can also be calculated using observable data.

Now define the expenditure shares for product Group 1 using just the products that are in Group 1. These are the shares that are relevant for the maximum overlap indexes which will be defined shortly. The *maximum overlap share* for product n in period t , s_{1n}^{tO} , is defined as follows:

$$s_{1n}^{tO} \equiv p_{1n}^t q_{1n}^t / p_1^t \cdot q_1^t; \quad t = 0, 1; n = 1, \dots, N. \quad (7)$$

These maximum overlap shares are also observable. It can be seen that the following

¹²The inner product of the vectors p_1^t and q_1^t is denoted as $p_1^t \cdot q_1^t \equiv \sum_{n=1}^N p_{1n}^t \cdot q_{1n}^t$, etc.

relationships hold between the true Group 1 shares and the maximum overlap Group 1 shares:¹³

$$\begin{aligned} s_{1n}^0 &= s_{1nO}^0 p_1^0 \cdot q_1^0 / [p_1^0 \cdot q_1^0 + p_3^0 \cdot q_3^0]; \quad n = 1, \dots, N; \\ &= s_{1nO}^0 \left[1 - \sum_{m=1}^M s_{3m}^0 \right]; \end{aligned} \quad (8)$$

$$\begin{aligned} s_{1n}^1 &= s_{1nO}^1 p_1^1 \cdot q_1^1 / [p_1^1 \cdot q_1^1 + p_2^0 \cdot q_2^0]; \quad n = 1, \dots, N; \\ &= s_{1nO}^1 \left[1 - \sum_{k=1}^K s_{2k}^1 \right]. \end{aligned} \quad (9)$$

3 The Törnqvist Price Index Decomposition

The US Bureau of Labor Statistics uses the Törnqvist price index as its target index for its chained Consumer Price Index (CPI).¹⁴ Typically, there are no adjustments for new and disappearing products so these Törnqvist price indexes are essentially maximum overlap price indexes. Let P_{TO} denote the *Törnqvist maximum overlap index*. This index is defined as follows:

$$P_{TO} \equiv \prod_{n=1}^N (p_{1n}^1 / p_{1n}^0)^{\frac{1}{2}(s_{1nO}^0 + s_{1nO}^1)}. \quad (10)$$

The *true Törnqvist index*, P_T , is defined as follows:

$$\begin{aligned} P_T &\equiv \prod_{n=1}^N \left(\frac{p_{1n}^1}{p_{1n}^0} \right)^{\frac{1}{2}(s_{1n}^0 + s_{1n}^1)} \prod_{k=1}^K \left(\frac{p_{2k}^1}{p_{2k}^{0*}} \right)^{\frac{1}{2}(s_{2k}^0 + s_{2k}^1)} \prod_{m=1}^M \left(\frac{p_{3m}^{1*}}{p_{3m}^0} \right)^{\frac{1}{2}(s_{3m}^0 + s_{3m}^1)} \\ &= \prod_{n=1}^N \left(\frac{p_{1n}^1}{p_{1n}^0} \right)^{\frac{1}{2}(s_{1n}^0 + s_{1n}^1)} \prod_{k=1}^K \left(\frac{p_{2k}^1}{p_{2k}^{0*}} \right)^{\frac{1}{2}(0 + s_{2k}^1)} \prod_{m=1}^M \left(\frac{p_{3m}^{1*}}{p_{3m}^0} \right)^{\frac{1}{2}(s_{3m}^0 + 0)} \\ &= \prod_{n=1}^N \left(\frac{p_{1n}^1}{p_{1n}^0} \right)^{\frac{1}{2}\{s_{1nO}^0 [1 - \sum_{m=1}^M s_{3m}^0] + s_{1nO}^1 [1 - \sum_{k=1}^K s_{2k}^1]\}} \prod_{k=1}^K \left(\frac{p_{2k}^1}{p_{2k}^{0*}} \right)^{\frac{1}{2}s_{2k}^1} \prod_{m=1}^M \left(\frac{p_{3m}^{1*}}{p_{3m}^0} \right)^{\frac{1}{2}s_{3m}^0} \\ &= P_{TO} \prod_{k=1}^K \left[\frac{(p_{2k}^1 / p_{2k}^{0*})}{\prod_{n=1}^N (p_{1n}^1 / p_{1n}^0)^{s_{1nO}^1}} \right]^{\frac{1}{2}s_{2k}^1} \prod_{m=1}^M \left[\frac{(p_{3m}^{1*} / p_{3m}^0)}{\prod_{n=1}^N (p_{1n}^1 / p_{1n}^0)^{s_{1nO}^0}} \right]^{\frac{1}{2}s_{3m}^0} \\ &= P_{TO} \times (1 + \mu) / (1 + \kappa) \end{aligned} \quad (11)$$

¹³These relationships are due to de Haan and Krsinich (2012; 31-32).

¹⁴See Bureau of Labor Statistics (2007).

where the second line uses (3) and (6), the fourth line uses (8) and (9) and the fifth line uses (10).¹⁵ The terms $1 + \kappa$ and $1 + \mu$ are defined as:

$$\begin{aligned}
1 + \kappa &\equiv \prod_{k=1}^K \left[\frac{p_{2k}^{0*}}{p_{2k}^1 / \prod_{n=1}^N (p_{1n}^1 / p_{1n}^0)^{s_{1nO}^1}} \right]^{\frac{1}{2} s_{2k}^1} \\
&= \prod_{k=1}^K \left[\frac{p_{2k}^{0*}}{p_{2k}^1 / P_{JO}^1} \right]^{\frac{1}{2} s_{2k}^1} \\
&= \prod_{k=1}^K \left[\frac{p_{2k}^{0*}}{p_{2kb}^0} \right]^{\frac{1}{2} s_{2k}^1} \tag{12}
\end{aligned}$$

$$\begin{aligned}
1 + \mu &\equiv \prod_{m=1}^M \left[\frac{p_{3m}^{1*}}{p_{3m}^0 \prod_{n=1}^N (p_{1n}^1 / p_{1n}^0)^{s_{1nO}^1}} \right]^{\frac{1}{2} s_{3m}^0} \\
&= \prod_{m=1}^M \left[\frac{p_{3m}^{1*}}{p_{3m}^0 P_{JO}^0} \right]^{\frac{1}{2} s_{3m}^0} \\
&= \prod_{m=1}^M \left[\frac{p_{3m}^{1*}}{p_{3mf}^1} \right]^{\frac{1}{2} s_{3m}^0} \tag{13}
\end{aligned}$$

where the (weighted) *Jevons index* using the maximum overlap share weights of period 1 is P_{JO}^1 and the (weighted) *Jevons index* using the maximum overlap share weights of period 0 is P_{JO}^0 ; i.e., these two indexes are defined as follows:¹⁶

$$P_{JO}^1 \equiv \prod_{n=1}^N \left(\frac{p_{1n}^1}{p_{1n}^0} \right)^{s_{1nO}^1}; \tag{14}$$

$$P_{JO}^0 \equiv \prod_{n=1}^N \left(\frac{p_{1n}^1}{p_{1n}^0} \right)^{s_{1nO}^0}, \tag{15}$$

and

$$p_{2kb}^0 \equiv p_{2k}^1 / P_{JO}^1; \quad k = 1, \dots, K; \tag{16}$$

$$p_{3mf}^1 \equiv p_{3m}^0 P_{JO}^0; \quad m = 1, \dots, M. \tag{17}$$

where p_{2kb}^0 are *carry backward prices* for the missing products in period 0 and the following and p_{3mf}^1 are *carry forward prices* for the missing products in period 1.

Equation (11) gives the relationship between the “true” cost of living index P_T and

¹⁵This formula was first derived by de Haan and Krsinich (2012; 31-32) (2014; 344). Their imputed prices for the missing products were obtained by using hedonic regressions whereas our imputed prices are interpreted as Hicksian reservation prices but the algebra is the same in both contexts. For additional discussion on this formula and its variants, see de Haan (2017).

¹⁶These could also be described as Cobb Douglas indexes, and (14) has been called a geometric Paasche index and (15) has been called a geometric Laspeyres index.

the price index P_{TO} that is defined over products that are available in both periods. The terms $1 + \kappa$ and $1 + \mu$ are counterparts to Feenstra's (1994; 159) λ_{t-1} and λ_t terms that he derived for bias due to changes in the availability of commodities in the context of CES preferences.

The inflation adjusted carry forward price defined by (17) for the missing product m in period 1 takes the observed price for product m in period 0, p_{3m}^0 and adjusts it for general inflation for the group of products that are present in both periods 0 and 1 using the (weighted) maximum overlap Jevons index P_{JO}^0 . Similarly, the inflation adjusted carry backward price defined by (16) for the missing product k in period 0 takes the observed price for product k in period 1, p_{2k}^1 and deflates it by the (weighted) Jevons maximum overlap price index, P_{JO}^1 . These inflation adjusted imputed prices are more reasonable than the often used *constant carry forward prices*, p_{3m}^0 , or *constant carry backward prices*, p_{2k}^1 . From (12), (13) and (11), it can be seen that if the reservation prices are equal to their inflation adjusted carry forward prices (so that $p_{3m}^{1*} = p_{3mf}^1$ for $m = 1, \dots, M$) and inflation adjusted carry backward prices (so that $p_{2k}^{0*} = p_{2kb}^0$ for $k = 1, \dots, K$), then the true Törnqvist index P_T will equal its maximum overlap counterpart, P_{TO} .

From the definition in (12), the term $1/(1 + \kappa)$ in (11) can be regarded as a measure of the *reduction* in the true cost of living due to the introduction of new products. The period 0 imputed price for new product k , p_{2k}^{0*} , is likely to be higher than the actual price for new product k in period 1 adjusted for general inflation, $p_{2kb}^0 = p_{2k}^1/P_{JO}^1$, and thus $1/(1 + \kappa)$ is likely to be less than 1. The bigger is the share of new products in period 1, $\sum_{k=1}^K s_{2k}^1$, the more $(1/1 + \kappa)$ will be less than 1.

The inflation adjustment term $1 + \mu$ defined by (13) can be regarded as a measure of the *increase* in the true cost of living due to the disappearance of existing products. The period 1 imputed price for disappearing product m , p_{3m}^{1*} , is likely to be higher than the actual price for product m in period 0 adjusted for general inflation, $p_{3mf}^1 = p_{3m}^0 P_{JO}^0$, and thus is likely to be greater than 1. The bigger is the share of disappearing products in period 0, $\sum_{m=1}^M s_{3m}^0$, the more μ will be greater than 1.

Note that the contribution of each new and disappearing product to the change in the true cost of living can be measured using the decomposition that definitions (11)-(17) provide. We define the following *margin terms*, κ_k and μ_m , which express how much higher each reservation price is from its inflation adjusted carry forward or backward price counterpart:

$$1 + \kappa_k \equiv p_{2k}^{0*}/p_{2kb}^0; \quad k = 1, \dots, K; \quad (18)$$

$$1 + \mu_m \equiv p_{3m}^{1*}/p_{3mf}^1; \quad m = 1, \dots, M. \quad (19)$$

Recall that the reservation price for a missing product is the price which would induce a utility maximizing potential purchaser of the product to demand zero units of it, hence

we expect the price to be higher than p_{2kb}^0 or p_{3mf}^1 , which from (16) and (17) are each observed prices adjusted by a maximum overlap inflation index.

Now substitute definitions (16)-(19) into (11) and we obtain the following *exact relationship* between the true Törnqvist index P_T and its maximum overlap counterpart P_{TO} :

$$\frac{P_T}{P_{TO}} = \frac{1 + \mu}{1 + \kappa} = \frac{\prod_{m=1}^M (1 + \mu_m)^{\frac{1}{2}s_{3m}^0}}{\prod_{k=1}^K (1 + \kappa_k)^{\frac{1}{2}s_{2k}^1}} \quad (20)$$

Hence, $(1 + \mu_m)^{\frac{1}{2}s_{3m}^0}$ is the contribution of each disappearing good m to the change in the true cost of living, and similarly $1/(1 + \kappa_k)^{\frac{1}{2}s_{2k}^1}$ is the contribution of each new good k .

From equations (1) and (2), the period 0 and 1 value aggregates for the goods and services in the group of $N + K + M$ commodities under consideration, v^0 and v^1 , are defined as follows:

$$v^0 \equiv p_1^0 \cdot q_1^0 + p_3^0 \cdot q_3^0; \quad v^1 \equiv p_1^1 \cdot q_1^1 + p_2^1 \cdot q_2^1. \quad (21)$$

The “true” *implicit Törnqvist quantity index* Q_T is defined as the value ratio, v^1/v^0 , deflated by the “true” Törnqvist price index, P_T ; i.e., we have:

$$Q_T \equiv [v^1/v^0]/P_T. \quad (22)$$

Statistical agencies can use maximum overlap Törnqvist price indexes to deflate final demand aggregates in order to construct aggregate quantity or volume indexes. Thus in our context, the *maximum overlap Törnqvist quantity index*, Q_{TO} , is defined as follows:

$$Q_{TO} = [v^1/v^0]/P_{TO}. \quad (23)$$

The bias in Q_{TO} relative to its true counterpart Q_T can be measured by the ratio Q_T/Q_{TO} :

$$Q_T/Q_{TO} = P_{TO}/P_T, \quad (24)$$

where we have used definitions (22) and (23) to derive (24). An exact expression for Q_T/Q_{TO} can then be obtained from (20):

$$\frac{Q_T}{Q_{TO}} = \frac{\prod_{k=1}^K (1 + \kappa_k)^{\frac{1}{2}s_{2k}^1}}{\prod_{m=1}^M (1 + \mu_m)^{\frac{1}{2}s_{3m}^0}} \quad (25)$$

If there are no disappearing goods, the right hand side of (25) becomes $\prod_{k=1}^K (1 + \kappa_k)^{\frac{1}{2}s_{2k}^1}$. Subtracting one and multiplying by 100, this number is a measure of the downward bias in the maximum overlap Törnqvist quantity index for the value aggregate in percentage points. That is, (25) gives the downward bias in the quantity index, and therefore in welfare change, from ignoring new goods and services.

In the following section, we develop analogous bias formulae for price and quantity

aggregates that are constructed using maximum overlap Laspeyres, Paasche or Fisher indexes.

4 The Laspeyres, Paasche and Fisher Decompositions

Using the notation that was defined in section 2, define the *true Laspeyres price index*, P_L , as follows:¹⁷

$$\begin{aligned}
P_L &\equiv [p_1^1 \cdot q_1^0 + p_2^1 \cdot q_2^0 + p_3^{1*} \cdot q_3^0] / [p_1^0 \cdot q_1^0 + p_2^{0*} \cdot q_2^0 + p_3^0 \cdot q_3^0] \\
&= [p_1^1 \cdot q_1^0 + p_3^{1*} \cdot q_3^0] / [p_1^0 \cdot q_1^0 + p_3^0 \cdot q_3^0] \quad \text{using } q_2^0 = 0_K \\
&= \sum_{n=1}^N s_{1n}^0 (p_{1n}^1 / p_{1n}^0) + \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*} / p_{3m}^0).
\end{aligned} \tag{26}$$

Define the *maximum overlap Laspeyres price index* P_{LO} that is defined only over products that are present in both periods as follows:

$$\begin{aligned}
P_{LO} &\equiv p_1^1 \cdot q_1^0 / p_1^0 \cdot q_1^0 \\
&= \sum_{n=1}^N s_{1nO}^0 (p_{1n}^1 / p_{1n}^0)
\end{aligned} \tag{27}$$

where the maximum overlap shares s_{1nO}^0 are defined above by definitions (7).

Recall equations (8) which exhibited the relationship between the true share weights for continuing products, s_{1n}^0 , and the share weights for the commodities present in each period, s_{1nO}^0 . Using definitions (26) and (27) and equations (8), we can derive the following

¹⁷The Lowe index, which uses a fixed base that may not be either of the periods under consideration, is used in constructing the CPI in many countries. We do not explicitly consider this index here, but similar results as for the Laspeyres index can of course be derived.

relationship between P_L and P_{LO} :

$$\begin{aligned}
P_L &\equiv \sum_{n=1}^N s_{1n}^0 (p_{1n}^1/p_{1n}^0) + \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*}/p_{3m}^0) \\
&= \sum_{n=1}^N s_{1nO}^0 \left[1 - \sum_{m=1}^M s_{3m}^0 \right] (p_{1n}^1/p_{1n}^0) + \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*}/p_{3m}^0) \\
&= P_{LO} - \sum_{n=1}^N s_{1nO}^0 \left(\sum_{m=1}^M s_{3m}^0 \right) (p_{1n}^1/p_{1n}^0) + \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*}/p_{3m}^0) \\
&= P_{LO} + \sum_{m=1}^M s_{3m}^0 \left[(p_{3m}^{1*}/p_{3m}^0) - \sum_{n=1}^N s_{1nO}^0 (p_{1n}^1/p_{1n}^0) \right] \\
&= P_{LO} + \sum_{m=1}^M s_{3m}^0 [(p_{3m}^{1*}/p_{3m}^0) - P_{LO}] \\
&= P_{LO} \left(1 + \sum_{m=1}^M s_{3m}^0 [(p_{3m}^{1*}/p_{3m}^0 P_{LO}) - 1] \right) \\
&= P_{LO} (1 + \alpha)
\end{aligned} \tag{28}$$

where $\alpha = \sum_{m=1}^M s_{3m}^0 [(p_{3m}^{1*}/p_{3m}^0 P_{LO}) - 1]$ is a bias term which when multiplied by P_{LO} and added to the maximum overlap Laspeyres price index, P_{LO} , enables us to obtain the true Laspeyres price index, P_L . There is a presumption that this error term will be *positive* in which case P_{LO} has a downward bias and must be adjusted upward by this term in order to account for the effective increase in the price level which is due to the disappearance of the products in Group 3 during period 1.

The decomposition defined by (28) is also useful in the context of defining *imputed carry forward prices* for products that become unavailable in period 1. Thus define the carry forward prices for the disappearing products as follows:¹⁸

$$p_{3mf}^1 \equiv p_{3m}^0 P_{LO}; \quad m = 1, \dots, M. \tag{29}$$

Thus the period 0 price for disappearing product m , p_{3m}^0 , is adjusted for general group inflation using the maximum overlap Laspeyres price index P_{LO} and is carried forward to period 1. These carry forward prices are different from the carry backward prices in the previous section which used the Jevons index P_{JO}^0 as the adjusting index of general inflation.

Define the following *margin terms*, α_m , which express how much higher each reservation price is from its Laspeyres inflation adjusted carry forward price counterpart:

$$1 + \alpha_m \equiv p_{3m}^{1*}/p_{3mf}^1; \quad m = 1, \dots, M. \tag{30}$$

¹⁸These new carry forward prices use the maximum overlap deflator P_{LO} instead of the maximum overlap Jevons deflator P_{JO}^1 which was used in our earlier definition of carry forward prices in (17).

Now substitute definitions (30) into (28) and we obtain the following *exact relationship* between the true Laspeyres price index P_L and its maximum overlap counterpart P_{LO} :

$$\begin{aligned}
P_L &\equiv P_{LO} + \sum_{m=1}^M s_{3m}^0 [(p_{3m}^{1*}/p_{3m}^0) - P_{LO}] \\
&= P_{LO} (1 + \sum_{m=1}^M s_{3m}^0 [(1 + \alpha_m) - 1]) \\
&= P_{LO} (1 + \sum_{m=1}^M s_{3m}^0 \alpha_m). \tag{31}
\end{aligned}$$

Thus we have the following exact formula for (P_L/P_{LO}) :

$$P_L/P_{LO} = 1 + \sum_{m=1}^M s_{3m}^0 \alpha_m. \tag{32}$$

We expect the margins α_m to be positive in general. In this case, it can be seen that the maximum overlap Laspeyres price index will understate inflation as measured by the true Laspeyres price index, provided that there are disappearing products in period 1.

The “true” implicit quantity index that matches up with the true Laspeyres price index is the *Paasche quantity index* Q_P defined as follows:

$$Q_P \equiv [v^1/v^0]/P_L. \tag{33}$$

The *maximum overlap Paasche quantity index*, Q_{PO} , is the implicit quantity index that deflates the value ratio by the maximum overlap Laspeyres price index:

$$Q_{PO} \equiv [v^1/v^0]/P_{LO}. \tag{34}$$

The bias in Q_{PO} relative to its true counterpart Q_P can be measured by the ratio Q_P/Q_{PO} :

$$\begin{aligned}
Q_P/Q_{PO} &= P_{LO}/P_L \text{ using (33) and (34)} \\
&= \left(1 + \sum_{m=1}^M s_{3m}^0 \alpha_m \right)^{-1} \text{ using (30)}. \tag{35}
\end{aligned}$$

Thus the use of a maximum overlap Laspeyres price index leads to a resulting maximum overlap quantity index which will tend to *overstate volume growth* if there are disappearing products in period 1.

A similar analysis can be carried out for the Paasche price index. Define the *true*

Paasche price index, P_P , as follows:

$$\begin{aligned}
P_P &\equiv [p_1^1 \cdot q_1^1 + p_2^1 \cdot q_2^1 + p_3^{1*} \cdot q_3^1] / [p_1^0 \cdot q_1^1 + p_2^{0*} \cdot q_2^1 + p_3^0 \cdot q_3^1] \\
&= [p_1^1 \cdot q_1^1 + p_2^1 \cdot q_2^1] / [p_1^0 \cdot q_1^1 + p_2^{0*} \cdot q_2^1] \quad \text{using } q_3^1 = 0_M \\
&= \left[\sum_{n=1}^N s_{1n}^1 (p_{1n}^1 / p_{1n}^0)^{-1} + \sum_{k=1}^K s_{2k}^1 (p_{2k}^1 / p_{2k}^{0*})^{-1} \right]^{-1}. \tag{36}
\end{aligned}$$

Define the *maximum overlap Paasche price index* P_{PO} that is defined only over products that are present in both periods as follows:

$$\begin{aligned}
P_{PO} &\equiv p_1^1 \cdot q_1^1 / p_1^0 \cdot q_1^1 \\
&= \left[\sum_{n=1}^N s_{1nO}^1 (p_{1n}^1 / p_{1n}^0)^{-1} \right]^{-1} \tag{37}
\end{aligned}$$

where the maximum overlap shares s_{1nO}^1 were defined above by definitions (7).

Recall equations (9) which exhibited the relationship between the period 1 true share weights for continuing products, s_{1n}^1 , and the corresponding share weights for the commodities present in each period, s_{1n}^{1*} . Using definitions (36) and (37) and equations (9), we can derive the following relationship between P_P and P_{PO} :

$$\begin{aligned}
(P_P)^{-1} &\equiv \sum_{n=1}^N s_{1n}^1 (p_{1n}^1 / p_{1n}^0)^{-1} + \sum_{k=1}^K s_{2k}^1 (p_{2k}^1 / p_{2k}^{0*})^{-1} \\
&= \sum_{n=1}^N s_{1nO}^1 \left[1 - \sum_{k=1}^K s_{2k}^1 \right] (p_{1n}^1 / p_{1n}^0)^{-1} + \sum_{k=1}^K s_{2k}^1 (p_{2k}^1 / p_{2k}^{0*})^{-1} \\
&= (P_{PO})^{-1} - \sum_{n=1}^N s_{1nO}^{1*} \sum_{k=1}^K s_{2k}^1 (p_{1n}^1 / p_{1n}^0)^{-1} + \sum_{k=1}^K s_{2k}^1 (p_{2k}^1 / p_{2k}^{0*})^{-1} \\
&= (P_{PO})^{-1} + \sum_{k=1}^K s_{2k}^1 \left[(p_{2k}^1 / p_{2k}^{0*})^{-1} - \sum_{n=1}^N s_{1nO}^1 (p_{1n}^1 / p_{1n}^0)^{-1} \right] \\
&= (P_{PO})^{-1} + \sum_{k=1}^K s_{2k}^1 \left[(p_{2k}^1 / p_{2k}^{0*})^{-1} - (P_{PO})^{-1} \right] \\
&= (P_{PO})^{-1} \left\{ 1 + \sum_{k=1}^K s_{2k}^1 \left[(p_{2k}^1 / p_{2k}^{0*} P_{PO})^{-1} - 1 \right] \right\} \\
&= (P_{PO})^{-1} \left\{ 1 + \sum_{k=1}^K s_{2k}^1 \left[(p_{2k}^{0*} P_{PO} / p_{2k}^1) - 1 \right] \right\} \\
&= (P_{PO})^{-1} (1 + \beta) \tag{38}
\end{aligned}$$

where $\beta \equiv \sum_{k=1}^K s_{2k}^1 [(p_{2k}^{0*} P_{PO} / p_{2k}^1) - 1]$ is a bias term which allows us to adjust the reciprocal of the Paasche price index defined over continuing products, $1/P_{PO}$, so that the adjusted

index is equal to the reciprocal of the true Paasche price index, $1/P_P$. The relationship (42) can be rewritten as follows:

$$P_P = P_{PO}(1 + \beta)^{-1}. \quad (39)$$

There is a presumption that the error term β will be *positive* in which case P_{PO} has an upward bias relative to P_P and hence must be adjusted downward by dividing P_{PO} by $(1 + \beta)$. This adjustment term $1/(1 + \beta)$ accounts for the effective decrease in the true index P_P which is due to the appearance of the new products in Group 2 during period 1.

The decomposition defined by (38) is also useful in the context of defining *imputed carry backward prices* for products that are available in period 1 but not period 0. Thus define the carry backward prices for the new products as follows:¹⁹

$$p_{2kb}^0 \equiv p_{2k}^1/P_{PO}; \quad k = 1, \dots, K. \quad (40)$$

Thus the period 1 price for new product k , p_{2k}^1 , is adjusted for general group inflation using the maximum overlap Paasche price index P_{PO} and is carried backward to period 0.

Define the following *margin terms*, β_k , which express how much higher each reservation price is from its Paasche inflation adjusted carry backward price counterpart:

$$1 + \beta_k \equiv p_{2k}^{0*}/p_{2kb}^0; \quad k = 1, \dots, K. \quad (41)$$

Now substitute definitions (41) into (38) and we obtain the following *exact relationship* between the true Paasche price index P_P and its maximum overlap counterpart P_{PO} :

$$\begin{aligned} (P_P)^{-1} &= (P_{PO})^{-1} \left\{ 1 + \sum_{k=1}^K s_{2k}^1 [(p_{2k}^1/p_{2kb}^0 P_{PO})^1 - 1] \right\} \\ &= (P_{PO})^{-1} \left\{ 1 + \sum_{k=1}^K s_{2k}^1 [(p_{2k}^{0*} P_{PO}/p_{2k}^1) - 1] \right\} \\ &= (P_{PO})^{-1} \left\{ 1 + \sum_{k=1}^K s_{2k}^1 [(1 + \beta_k) - 1] \right\} \\ &= (P_{PO})^{-1} \left(1 + \sum_{k=1}^K s_{2k}^1 \beta_k \right). \end{aligned} \quad (42)$$

Thus we have the following exact formula for P_{PO}/P_P :

$$P_{PO}/P_P = 1 + \sum_{k=1}^K s_{2k}^1 \beta_k. \quad (43)$$

¹⁹These new carry backward prices use the maximum overlap deflator P_{PO} instead of the maximum overlap Jevons deflator P_{JO}^1 which was used in our earlier definition of carry backward prices in (16).

We expect the margins β_k to be positive in general. In this case, it can be seen that the maximum overlap Paasche price index P_{PO} will overstate inflation as measured by the true Paasche price index P_P , provided that there are new products in period 1.

Equation (32) gave an expression for (P_L/P_{LO}) . It is convenient to have a companion formula for (P_P/P_{PO}) . Rearranging (42) we have the following exact formula:

$$P_P/P_{PO} = \left(1 + \sum_{k=1}^K s_{2k}^1 \beta_k \right)^{-1} \quad (44)$$

The “true” implicit quantity index that matches up with the true Paasche price index is the *Laspeyres quantity index* Q_L defined as follows:

$$Q_L \equiv [v^1/v^0]/P_P. \quad (45)$$

The *maximum overlap Laspeyres quantity index*, Q_{LO} , is the implicit quantity index that deflates the value ratio by the maximum overlap Paasche price index:

$$Q_{LO} \equiv [v^1/v^0]/P_{PO}. \quad (46)$$

The bias in Q_{LO} relative to its true counterpart Q_L can be measured by the ratio Q_L/Q_{LO} :

$$\begin{aligned} Q_L/Q_{LO} &= P_{PO}/P_P \text{ using (45) and (46)} \\ &= 1 + \sum_{k=1}^K s_{2k}^1 \beta_k \text{ using (43)}. \end{aligned} \quad (47)$$

Thus the upward bias in the maximum overlap Paasche price index P_{PO} translates into a downward bias in the companion maximum overlap Laspeyres quantity index, Q_{LO} .

Finally, we look at the bias in the maximum overlap Fisher indexes relative to their true counterparts. Define the true Fisher index, P_F , as the geometric mean of the true Laspeyres and Paasche indexes and define the maximum overlap Fisher index over commodities present in both periods, P_{FO} , as the geometric mean of the maximum overlap Laspeyres and Paasche indexes, P_{LO} and P_{PO} :

$$P_F \equiv (P_L P_P)^{1/2}; \quad (48)$$

$$P_{FO} \equiv (P_{LO} P_{PO})^{1/2}. \quad (49)$$

The exact relationship between P_F and P_{FO} can be determined by substituting the exact

decompositions for P_L and P_P given by (28) and (39) into definition (48):

$$\begin{aligned}
P_F &\equiv (P_L P_P)^{1/2} \\
&= [P_{LO}(1 + \alpha)P_{PO}(1 + \beta)^{-1}]^{1/2} \\
&= P_{FO}[(1 + \alpha)(1 + \beta)^{-1}]^{1/2} \text{ using definition (49)}
\end{aligned} \tag{50}$$

The Fisher index decomposition defined by (50) is a counterpart to the Törnqvist approximate decomposition defined earlier by (20). Typically, they can be expected to give much the same answer. If either is the target index of interest, then using our framework it becomes easy to derive expressions for possible biases. For example, if the Fisher price index is the target index, then the bias in the Laspeyres maximum overlap index, P_{LO} , can be written as follows:

$$P_F - P_{LO} = [P_L P_P]^{1/2} - P_{LO} = P_{LO} \left\{ \left[\frac{P_{PO}(1 + \alpha)}{P_{LO}(1 + \beta)} \right]^{1/2} - 1 \right\} \tag{51}$$

We can note that the adjustment term for the downward bias in the Laspeyres maximum overlap index, $1 + \alpha$, is offset by the adjustment term for the upward bias in the Paasche maximum overlap index, $1 + \beta$. That is, while both adjustment terms can be expected to be greater than one, given that the entry rate of new products seems to typically exceed exit rates, it is likely that the upward bias associated with disappearing products will be more than offset by the downward bias associated with the new products, so that $(1 + \alpha)/(1 + \beta) > 1$.²⁰ The Paasche-Laspeyres spread of the maximum overlap price indexes, P_{PO}/P_{LO} , will likely be smaller than unity and an increasing spread will tend to increase the bias. Also, the level of the maximum overlap Laspeyres index, P_{LO} , will feed directly into the percentage point bias. Expressions for biases of other indexes relative to target indexes can easily be worked out in a similar fashion.

The “true” *implicit Fisher quantity index* Q_F is defined as the value ratio, v^1/v^0 , deflated by the “true” Fisher price index P_F ; i.e., we have:

$$Q_F \equiv [v^1/v^0]/P_F. \tag{52}$$

Some statistical agencies use maximum overlap Fisher price indexes to deflate final demand aggregates in order to construct aggregate quantity or volume indexes. Thus in our context, the *maximum overlap Fisher quantity index*, Q_{FO} , is defined as follows:

$$Q_{FO} \equiv [v^1/v^0]/P_{FO}. \tag{53}$$

²⁰If, for example, upward and downward biases arising from individual items are the same, and shares are the same, then the higher entry rate of new products will mean that the downward bias dominates. Of course, the reverse could also occur.

The bias in Q_{FO} relative to its true counterpart Q_F can be measured by the ratio Q_F/Q_{FO} :

$$Q_F/Q_{FO} = P_{FO}/P_F \quad (54)$$

where we have used definitions (52) and (53) to derive (54). An exact expression for Q_F/Q_{FO} can be obtained from (50):

$$\begin{aligned} Q_F/Q_{FO} &\equiv [(1 + \alpha)^{-1}(1 + \beta)]^{1/2} \\ &= \left[\frac{1 + \sum_{k=1}^K s_{2k}^1 \beta_k}{1 + \sum_{m=1}^M s_{3m}^0 \alpha_m} \right]^{1/2}, \end{aligned} \quad (55)$$

which is the geometric mean of the biases in the Paasche and Laspeyres cases from (35) and (47). If there are no disappearing goods, the right hand side of (55) becomes $\left(1 + \sum_{k=1}^K s_{2k}^1 \beta_k\right)^{1/2}$. Subtracting one and multiplying by 100, this number is a measure of the downward bias in the maximum overlap Fisher quantity index for the value aggregate in percentage points. That is, (55) gives the downward bias in the quantity index, and therefore in welfare change, from ignoring new goods and services.

There is a similarity between the expression for the bias in the Fisher index case in (55) and that for the Törnqvist index in (25). Subtract 1 from both sides of (25), define the right hand side of the resulting expression as the function $h(\kappa_1, \dots, \kappa_K, \mu_1, \dots, \mu_M)$ and approximate h by taking the first order Taylor series approximation to h evaluated at $0 = \kappa_1 = \dots = \kappa_K = \mu_1 = \dots = \mu_M$. The resulting approximation to $(Q_T/Q_{TO}) - 1$ is the following one:²¹

$$(Q_T/Q_{TO}) - 1 \approx \frac{1}{2} \left(\sum_{k=1}^K s_{2k}^1 \kappa_k - \sum_{m=1}^M s_{3m}^0 \mu_m \right). \quad (56)$$

Similarly, subtract 1 from both sides of (55), define the right hand side of the resulting expression as a function of $\alpha_1, \dots, \alpha_m$ and β_1, \dots, β_k and form the first order Taylor series approximation to this function around $0 = \kappa_1 = \dots = \kappa_K = \mu_1 = \dots = \mu_M$. The resulting approximation to $(Q_F/Q_{FO}) - 1$ is the following one:

$$(Q_F/Q_{FO}) - 1 \approx \frac{1}{2} \left(\sum_{k=1}^K s_{2k}^1 \beta_k - \sum_{m=1}^M s_{3m}^0 \alpha_m \right). \quad (57)$$

Recalling the definitions κ_k and β_k from (18) and (41), they are the same except that κ_k uses the (weighted) maximum overlap Jevons index to impute the carry backward prices while β_k uses the maximum overlap Paasche index. Similarly, the imputed carry forward prices in the definition of μ_m from (19) use the (weighted) maximum overlap Jevons index, while the α_k from (30) use the maximum overlap Laspeyres index. Hence, from (56) and

²¹This formula is similar in spirit to the highly simplified approximate formulae obtained by Diewert (1987; 779) (1998; 51-54).

(57), it can be expected that the biases from both indexes will be empirically quite similar.

In the next section, we adapt the algebra developed in this section to the problems associated with replacing disappearing products with closely related substitute products.

5 Biases from Replacement Sampling with Quality Change

Triplett (2004; 12-40) provides an excellent discussion of statistical agency practices for dealing with quality change in the context of *replacement sampling*. In most elementary strata, products disappear from a sample of similar products while new products appear. It is common for statistical agencies to refresh their sample of products by substituting replacement products for the disappearing products. However, in order to make the replacement products comparable to the disappearing products, the statistical agency may make some *quality adjustments* to the new products. Triplett systematically describes the main methods of quality adjustment used by statistical agencies in his Handbook. In this section we will discuss most of methods he reviewed.²²

We will study the possible biases in sample replacement methods in the transaction data (e.g. scanner data) context; i.e., we will assume that price and quantity data are available to the statistical agency for the set of products under consideration.²³ We will also adapt our Hicksian reservation price methodology that was described in section 4 to the product replacement context; i.e., we will continue to assume that M products disappear in period 1 but in this section, we will assume that the K new products introduced in period 1 are replacement products for the M disappearing products (so that now $K = M$). We will adapt the algebra in section 4 first to the case where the Laspeyres index is the target index. Subsequently, we will consider the companion case where the Paasche index is the target index. The case of the Fisher index as the target is left to the reader.

As in the previous sections, all the biases identified for these price indexes will result in biases in the other direction in the corresponding quantity indexes. Hence, the results highlight the potential for mismeasurement arising from quality adjustment methods, not

²²Triplett's (2004) Handbook is mostly about the use of hedonic regression models to perform quality adjustment on new products to make them comparable to disappearing products. However, he realized that there was considerable opposition to the use of hedonic regression models by official statistical agencies due to their "subjective" nature. Triplett probably wrote his Chapter II as a response to these criticisms; i.e., he showed that existing statistical agency practices in dealing with disappearing products also had large subjective aspects!

²³Hence, we do not consider the case of free goods, but focus on transacted goods, as per statistical agency practice. It can be argued that with transaction level data, replacement strategies are unnecessary: "Given some index number formula, quality adjustment is a matter of imputation or prediction of 'missing' prices or price relatives. What is needed is an estimate of what the price or price relative of a disappearing (new) item would have been, had it been sold during the current (base) period. With prediction comes statistical modelling, which is usually hedonic modelling in this context." Jan de Haan (2007; 1). However, many statistical agencies still use a traditional sample replacement approach.

only directly in the price indexes, but also in the corresponding measures of interest such as real consumption, real output and productivity.

5.1 The Laspeyres Case

We now make the assumption that the new products in Group 2 that appear in period 1 are *replacement products* for the M disappearing products in period 1. Thus we set $K = M$ and the p_{2k}^1 which appear in sections 2-4 now become the M observable replacement prices in Group 2, which we label as p_{2m}^1 for $m = 1, \dots, M$.

However, the statistical agency may quality adjust these replacement prices to make them more comparable to the period 0 product prices for Group 3, the p_{3m}^0 . The quality adjustment factor for product m is A_m^1 for $m = 1, \dots, M$. Thus the statistical agency *quality adjusted replacement price* for disappearing product m in period 1 is p_{3mr}^1 defined as the period 1 observable Group 2 (replacement) product price p_{2m}^1 times the quality adjustment factor A_m^1 :

$$p_{3mr}^1 \equiv A_m^1 p_{2m}^1; \quad m = 1, \dots, M. \quad (58)$$

We define the *replacement Laspeyres index*, P_{LR} , in the same way as we defined the true Laspeyres index P_L as in (26) except that the true reservation prices, p_{3m}^{1*} , are replaced by the quality adjusted replacement prices p_{3mr}^1 defined by (58). Thus we have the following definition:

$$\begin{aligned} P_{LR} &\equiv \sum_{n=1}^N s_{1n}^0 (p_{1n}^1 / p_{1n}^0) + \sum_{m=1}^M s_{3m}^0 (p_{3mr}^1 / p_{3m}^0) \\ &= \sum_{n=1}^N s_{1n}^0 (p_{1n}^1 / p_{1n}^0) + \sum_{m=1}^M s_{3m}^0 (A_m^1 p_{2m}^1 / p_{3m}^0). \end{aligned} \quad (59)$$

We can form an estimate for the *bias* in P_{LR} by taking the difference between the true Laspeyres index P_L defined by (26) and (59):

$$\begin{aligned} P_L - P_{LR} &= \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*} / p_{3m}^0) - \sum_{m=1}^M s_{3m}^0 (p_{3mr}^1 / p_{3m}^0) \\ &= \sum_{m=1}^M s_{3m}^0 (p_{3m}^0)^{-1} (p_{3m}^{1*} - A_m^1 p_{2m}^1). \end{aligned} \quad (60)$$

Thus if $A_m^1 \geq p_{3m}^{1*} / p_{2m}^1$ for $m = 1, \dots, M$ with at least one strict inequality (so that the quality adjustment factors A_m^1 are *too large*), then P_{LR} will be less than the true Laspeyres index and hence will have a *downward bias*. If $A_m^1 \leq p_{3m}^{1*} / p_{2m}^1$ for $m = 1, \dots, M$ with at least one strict inequality (so that the quality adjustment factors A_m^1 are *too small*), then P_{LR} will be greater than the true Laspeyres index and hence will have an *upward bias*.

Now we can analyze some of Triplett's (2004; 21-29) special cases.²⁴ These cases are examples of actual practice by national statistical offices, in both in the Consumer Price Index and Producer Price Index; see e.g. Bureau of Labor Statistics (2014), Wells and Restieaux (2014), and Adams and Klayman (2018).

5.1.1 Special Case 1: Triplett's Direct Comparison Method

This is the special case of (60) where the quality adjustment factors are all chosen to equal 1; i.e., we have $A_m^1 = 1$ for $m = 1, \dots, M$. Thus the period 1 replacement product prices p_{2m}^1 are regarded as exact substitute prices for the corresponding disappearing product prices p_{3m}^1 (which are missing). Substitute the equations $A_m^1 = 1$ into (60) above in order to obtain the following bias formula for the *replacement Laspeyres index that uses the direct comparison method*:

$$P_L - P_{LR} = \sum_{m=1}^M s_{3m}^0 (p_{3m}^0)^{-1} [p_{3m}^{1*} - p_{2m}^1]. \quad (61)$$

If product $2m$ is a perfect substitute for product $3m$ for $m = 1, \dots, M$, then p_{3m}^{1*} will equal p_{2m}^1 for all m and P_{LR} will equal the true Laspeyres price index, P_L . If product $2m$ has approximately the same quality as product $3m$ for each m but the product pairs are not perfect substitutes, then the reservation prices p_{3m}^{1*} will tend to be higher than the corresponding prices p_{2m}^1 (these products are actually available in period 1 and hence should be less than their reservation prices) and thus the right hand side of (61) will tend to be positive. Hence P_{LR} will tend to have a downward bias relative to the true Laspeyres index P_L .²⁵

5.1.2 Special Case 2: Triplett's Link to Show No Change Case (Carry Forward Method)

Perhaps a better description of this method for quality adjustment would be to call this method *the price carry forward with no inflation adjustment method*; i.e., we simply assume that the period 1 replacement price for product m is the corresponding base period price p_{3m}^0 . In this case, we have $p_{3mr}^1 \equiv p_{3m}^0 = A_m^1 p_{2m}^1$ for $m = 1, \dots, M$. Hence, the quality adjustment factors are $A_m^1 = p_{3m}^0 / p_{2m}^1$. Using (60), the resulting bias formula is now the

²⁴Other cases can be considered using the same framework.

²⁵Thus this downward replacement bias will tend to offset some of the upward substitution bias in the true Laspeyres index.

following one:

$$\begin{aligned}
P_L - P_{LR} &= \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*}/p_{3m}^0) - \sum_{m=1}^M s_{3m}^0 (p_{3mr}^1/p_{3m}^0) \\
&= \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*}/p_{3m}^0) - \sum_{m=1}^M s_{3m}^0 (p_{3m}^0/p_{3m}^0) \quad \text{using } p_{3mr}^1 = p_{3m}^0 \\
&= \sum_{m=1}^M s_{3m}^0 [(p_{3m}^{1*}/p_{3m}^0) - 1]. \tag{62}
\end{aligned}$$

If there is general inflation for the commodity group under consideration going from period 0 to 1, then P_{LR} will tend to have a downward bias due to the fact that the period 0 prices p_{3m}^0 are not adjusted upwards for inflation. There may be another dose of downward bias in P_{LR} due to the fact that the reservation prices p_{3m}^{1*} will tend to be higher than the prices of substitute products that are available in period 1. If there is general deflation for the commodity group under consideration going from period 0 to 1, then P_{LR} will tend to have an upward bias due to the fact that the period 0 prices p_{3m}^0 upward bias could be offset by the fact that the reservation prices p_{3m}^{1*} will tend to be higher than the prices of substitute products that are available in period 1. In general, this method is not recommended.

5.1.3 Special Case 3: Triplett's Deletion Method (Maximum Overlap Method)

Perhaps a better description of this method for quality adjustment would be to call this method *the price carry forward with a maximum overlap inflation adjustment method*. In this case, we assume that the period 1 replacement price for missing product m is the corresponding base period price p_{3m}^0 times an index of general inflation. The quality adjustment factors are then $A_m^1 \equiv P_{LO} p_{3m}^0 / p_{2m}^1$. The index of general inflation that we use is the maximum overlap Laspeyres price index, P_{LO} , defined by (27) in section 4.²⁶ Thus the replacement prices for the missing M products in period 1 are defined as follows for this method:

$$\begin{aligned}
p_{3mr}^1 &\equiv p_{3m}^0 P_{LO}; \quad m = 1, \dots, M \\
&= p_{3mf}^1 \tag{63}
\end{aligned}$$

where the second set of equalities in (63) follows from definitions (29); i.e., the p_{3mf}^1 are the *Laspeyres maximum overlap carry forward prices* for the disappearing products that

²⁶Aghion, Bergeaud, Boppart, Klenow and Li (2017) note that using the price change of continuing products as the inflation rate for new varieties leads to an underestimation of real growth if the replacement products are of higher quality. Their method uses strong assumptions to identify the evolution of the market share for continuing products and in the translation to a cost of living index. Such assumptions are not used here.

were defined in section 4.

Substitute equations (63) into equation (60) and we obtain the following equations:²⁷

$$\begin{aligned}
P_L - P_{LR} &= \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*}/p_{3m}^0) - \sum_{m=1}^M s_{3m}^0 (p_{3mr}^1/p_{3m}^0) \\
&= \sum_{m=1}^M s_{3m}^0 (p_{3m}^{1*}/p_{3m}^0) - \sum_{m=1}^M s_{3m}^0 (P_{LO}) \quad \text{using (63)} \\
&= P_{LO} \sum_{m=1}^M s_{3m}^0 [(p_{3m}^{1*}/P_{LO}p_{3m}^0) - 1] \\
&= P_{LO} \sum_{m=1}^M s_{3m}^0 [(p_{3m}^{1*}/p_{3mf}^1) - 1] \quad \text{using (63)}. \tag{64}
\end{aligned}$$

Thus the difference between the true Laspeyres index P_L and the Laspeyres index using inflation adjusted carry forward prices p_{3mf}^1 as approximations to the Hicksian reservation prices p_{3m}^{1*} hinges on the differences between these two sets of prices. In general, we would expect the reservation prices p_{3m}^{1*} to be greater than their inflation adjusted carry forward price counterparts p_{3mf}^1 . Thus in general, we expect P_{LR} to have a downward bias using this method.²⁸ However, if general inflation is positive, then the bias using this method should be considerably less than the bias using Method 2 above.

5.2 The Paasche Case

We now consider how the replacement methodology used above works in the context of evaluating a true Paasche index and various approximations to it. The true Paasche index, P_P , was defined by (36) in section 4. However, in the present context, there are M new products instead of K new products. These M new products are Group 2 products that replace the Group 3 products that disappeared in period 1. For convenience, we write the reciprocal of the true Paasche index in the present context as follows:

$$(P_P)^{-1} = \sum_{n=1}^N s_{1n}^1 (p_{1n}^1/p_{1n}^0)^{-1} + \sum_{m=1}^M s_{2m}^1 (p_{2m}^1/p_{2m}^{0*})^{-1}. \tag{65}$$

The prices for the new products which have appeared in period 1 as replacement products are now the prices p_{2m}^1 for $m = 1, \dots, M$ and we require *reservation prices* for these products in period 0 which we now denote as p_{2m}^{0*} for $m = 1, \dots, M$. If these replacement products

²⁷In this case, P_{LR} coincides with P_{LO} . The bias formula (7) is a generalization of a bias formula due to Triplett (2004; 25).

²⁸This general verdict on the method agrees with Triplett's (2004; 29) general assessment of the bias in this method. However, Triplett goes on to state that in the case of computers or other products that have rapid downward price change, the bias may go in the other direction. Thus the direction of bias will depend on the context.

were not available in period 0, then the p_{2m}^{0*} are true Hicksian reservation prices. However, it may be the case that products m existed in period 0 in which case, the period 0 observed price for this product, p_{2m}^0 , could be observed. In this case, the reservation price should be taken to be the observed price; i.e., in this case, we have:

$$p_{2m}^{0*} \equiv p_{2m}^0 \quad \text{if product } m \text{ exists in period 0.} \quad (66)$$

In the general case where either the M replacement products did not exist in period 0 or where the statistical agency is unable to collect observed period 0 prices for these products, then in order to construct an approximation to the true Paasche price index, the statistical agency may insert *estimated prices* or *replacement prices* p_{2mr}^0 in place of the “true” period 0 reservation prices p_{2m}^{0*} for $m = 1, \dots, M$. In order to make the analysis of the Paasche replacement index symmetric to our treatment of the Laspeyres replacement index, we introduce a new set of *product specific inflation adjustment factors*, A_m^0 that convert the observed period 1 prices p_{2m}^1 into their estimated counterpart prices that approximate the period 0 reservation prices p_{2m}^{0*} .²⁹

$$p_{2mr}^0 \equiv p_{2m}^1 / A_m^0; \quad m = 1, \dots, M. \quad (67)$$

Now define *the replacement Paasche index*, P_{PR} , in the same way as one would define the true Paasche index P_P as in (65) except that the true period 0 reservation prices, p_{2m}^{0*} , are replaced by the inflation adjusted replacement prices p_{2mr}^0 defined by (67). Thus we have the following definition for the reciprocal of P_{PR} :

$$\begin{aligned} (P_{PR})^{-1} &= \sum_{n=1}^N s_{1n}^1 (p_{1n}^1 / p_{1n}^0)^{-1} + \sum_{m=1}^M s_{2m}^1 (p_{2m}^1 / p_{2mr}^0)^{-1} \\ &= \sum_{n=1}^N s_{1n}^1 (p_{1n}^1 / p_{1n}^0)^{-1} + \sum_{m=1}^M s_{2m}^1 (A_m^0)^{-1} \quad \text{using (67).} \end{aligned} \quad (68)$$

We can form an estimate for the *bias* in P_{LR} by taking the difference between (65) and (68):

$$\begin{aligned} (P_P)^{-1} - (P_{PR})^{-1} &= \sum_{m=1}^M s_{2m}^1 (p_{2m}^1 / p_{2m}^{0*})^{-1} - \sum_{m=1}^M s_{2m}^1 (A_m^0)^{-1} \\ &= \sum_{m=1}^M s_{2m}^1 [(p_{2m}^1 / p_{2m}^{0*})^{-1} - (A_m^0)^{-1}]. \end{aligned} \quad (69)$$

Thus if $A_m^0 \geq p_{2m}^1 / p_{2m}^{0*}$ for $m = 1, \dots, M$ with at least one strict inequality (so that the

²⁹Quality adjustment is not an issue in defining the “quality” adjustment factors A_m^0 if all of the replacement products actually existed in period 0. Thus we will refer to the A_m^0 as inflation adjustment factors.

product specific inflation adjustment factors A_m^0 are *too large*), then $(P_{PR})^{-1}$ will be less than the reciprocal of the true Paasche index. Thus P_{PR} will be greater than the true Paasche index P_P , and hence will have an *upward bias*. If $A_m^0 \leq p_{2m}^1/p_{2m}^{0*}$ for $m = 1, \dots, M$ with at least one strict inequality (so that the product specific inflation adjustment factors A_m^0 are *too small*), then P_{PR} will be less than the true Paasche index P_P , and hence will have a *downward bias*.

We will consider three special cases of the above general methodology.³⁰

5.2.1 Special Case 1: The Use of Inflation Adjusted Paasche Carry Backward Prices

In this case, the replacement prices for period 0 are set equal to the inflation adjusted carry backward prices defined by equations (40). Thus using our present notation, we have the following replacement prices:

$$p_{2mr}^0 \equiv p_{2m}^1/A_m^0 = p_{2mb}^0 \equiv p_{2m}^1/P_{PO}; \quad m = 1, \dots, M \quad (70)$$

where P_{PO} is the Paasche maximum overlap price index defined by (37), which in this case is equal to the Paasche replacement price index defined by (68) where the replacement prices p_{2mr}^0 are defined by equations (70). Note that in this case, all of the product specific inflation factors A_m^0 are equal to the maximum overlap Paasche index, P_{PO} , which in turn is also equal to P_{PR} in this case. Thus the bias in this case can be determined by evaluating the following version of equation (69) where the A_m^0 are all equal to P_{PO} :

$$\begin{aligned} (P_P)^{-1} - (P_{PR})^{-1} &= \sum_{m=1}^M s_{2m}^1 [(p_{2m}^1/p_{2m}^{0*})^{-1} - (A_m^0)^{-1}] \\ &= \sum_{m=1}^M s_{2m}^1 [(p_{2m}^1/p_{2m}^{0*})^{-1} - (P_{PO})^{-1}]. \end{aligned} \quad (71)$$

If the Group 2 products were missing in period 0, then the reservation prices p_{2m}^{0*} are likely to be relatively high and the price ratios p_{2m}^1/p_{2m}^{0*} are likely to be less than the maximum overlap Paasche index P_{PO} , so that we are likely to have $(p_{2m}^1/p_{2m}^{0*})^{-1} > (P_{PO})^{-1}$ so the right hand side of (71) is likely to be positive. Thus $(P_P)^{-1} > (P_{PR})^{-1}$ which in turn implies $P_P < P_{PR} = P_{PO}$. Thus in this case, P_{PR} is likely to have an upward bias.

5.2.2 Special Case 2: The Case of Carry Backward Prices with No Inflation Adjustment

This is the Paasche counterpart to Special Case 2 for the Laspeyres index. In this case, the replacement prices for period 0 are set equal to the period 1 product prices for the

³⁰Other cases can be considered using the same framework.

replacement products. Thus we have the following replacement prices:

$$p_{2mr}^0 \equiv p_{2m}^1/A_m^0 \equiv p_{2m}^1; \quad m = 1, \dots, M. \quad (72)$$

In this case, $A_m^0 = 1$ for all m . The bias in this case can be determined by evaluating the following version of equation (69) where the A_m^0 are all equal to 1:

$$\begin{aligned} (P_P)^{-1} - (P_{PR})^{-1} &= \sum_{m=1}^M s_{2m}^1 [(p_{2m}^1/p_{2m}^{0*})^{-1} - (A_m^0)^{-1}] \\ &= \sum_{m=1}^M s_{2m}^1 [(p_{2m}^1/p_{2m}^{0*})^{-1} - (1)^{-1}]. \end{aligned} \quad (73)$$

If there is a large amount of general inflation or deflation between periods 0 and 1 for the product category under consideration, then we are probably safe in asserting that the bias in using this method will be greater than the bias using the method described in the first special case (because the present method does not account for general inflation between periods 0 and 1).

5.2.3 Special Case 3: Prices for the Replacement Products are Available in Period 0

In this case, we assume that the statistical agency can observe these period 0 prices, p_{2m}^0 for $m = 1, \dots, M$. The reservation prices p_{2m}^{0*} are then equal to the observed prices, p_{2m}^0 , for $m = 1, \dots, M$. The product specific inflation adjustment factors A_m^0 in this case are the observed price ratios for the products:

$$A_m^0 \equiv p_{2m}^1/p_{2m}^0; \quad m = 1, \dots, M. \quad (74)$$

In this case, the reciprocal of the replacement Paasche index, P_{PR} , is defined as follows:

$$\begin{aligned} (P_{PR})^{-1} &= \sum_{n=1}^N s_{1n}^1 (p_{1n}^1/p_{1n}^0)^{-1} + \sum_{m=1}^M s_{2m}^1 (A_m^0)^{-1} \\ &= \sum_{n=1}^N s_{1n}^1 (p_{1n}^1/p_{1n}^0)^{-1} + \sum_{m=1}^M s_{2m}^1 (p_{2m}^1/p_{2m}^0)^{-1} \\ &\equiv P_P^{-1}. \end{aligned} \quad (75)$$

Thus under these conditions, the replacement Paasche index is equal to the true Paasche index. Hence, if possible, the statistical agency should attempt to get transactions prices in the previous period for the replacement products it uses in its index computations for the current period. Of course, this will usually not be possible.

It can be seen that Triplett was quite right in flagging the problem of sample attrition

and replacement methodologies as a serious problem with traditional statistical agency procedures.

6 Conclusion

The problem of accounting for new and disappearing goods is not a new one. However, concerns about potential mismeasurement have heightened with the observation of high product turnover, accompanied by rapid quality change and the introduction of transformative innovations arising from the Digital Economy, at a time when industrialized countries are experiencing a prolonged productivity slowdown. To aid understanding of the potential for mismeasurement to be playing a role in influencing key economic indicators, such as the CPI, real GDP and productivity growth, we derived exact expressions for the biases in price and quantity indexes that arise from the use of standard statistical agency index number methodology and quality adjustment methods.

Statistical agencies almost always construct estimates of a real variable by deflating the value aggregate of interest by using a particular index number formula. However, in the context of new and disappearing products instead of using the “true” index number formula for the price index, statistical agencies often use a maximum overlap index that is restricted to the products present in both periods being compared.

We evaluated the bias in the use of maximum overlap indexes for the Törnqvist, Laspeyres, Paasche and Fisher price and quantity indexes in sections 3 and 4. These are of interest, as the Törnqvist index is the target index for the CPI in the U.S. and is used to construct aggregates for productivity in e.g. the U.S. (outputs and inputs) and Australia (inputs),³¹ the Laspeyres index is used by most countries in constructing real GDP and the CPI,³² the GDP deflator will have the Paasche index form if the target real GDP index is Laspeyres, and the Fisher index is used to construct U.S. real GDP.

The resulting bias formulae are very simple but require estimates of Hicksian reservation prices for the missing products. These reservation prices may not be easy to compute, and in some cases they may be out of scope for official price indices, at least currently.³³ Also, calculations of the biases presented here are typically not producible on the timescale of official statistics, at least currently. Indeed, if the biases could be calculated in a timely fashion, then there would not be an issue – the true indexes could be calculated directly, or the current indexes adjusted for biases using the formulae we have provided.

We have also evaluated biases which arise from the practice of sample replacement due

³¹Following Ivancic, Diewert and Fox (2011), de Haan and van der Grient (2011) and Diewert and Fox (2017b), from the December quarter 2017 the Australian Bureau of Statistics switched to using a Törnqvist-based multilateral index for incorporating transaction data into the CPI; see ABS (2017).

³²More correctly, for the CPI the Lowe index is typically used, of which the Laspeyres formula is a special case; see Diewert (1993a).

³³See Reinsdorf and Schreyer (2017).

to disappearing goods, with the replacement goods possibly being of a different quality. Specifically, explicit expressions for the biases arising from different quality adjustment methods, which are common candidates for use by statistical agencies in practice, have been provided for standard index number formulae.

Such expressions inform practitioners and policy makers of the biases inherent in different methodologies. They provide a theoretical basis and framework for the emerging empirical literature on new goods and services, and to quality adjustment methodologies used in practice.

In modern economies with rapid product turnover, with a proliferation of digital products and services and quality change, it is timely to finally have clarity about the biases in standard methodologies as identified in this paper.

References

- ABS (2017), “An Implementation Plan to Maximise the Use of Transactions Data in the CPI”, Australian Bureau of Statistics Information Paper, Catalogue No. 6401.0.60.004, Canberra.
- Adams, B. and A. Klayman (2018), “Cross-validation of Quality-adjustment Methods for Price Indexes”, Monthly Labor Review, June, U.S. Bureau of Labor Statistics, Washington, D.C. <https://doi.org/10.21916/mlr.2018.18>.
- Aghion P., A. Bergeaud, T. Boppart, P.J. Klenow and H. Li (2017), “Missing Growth from Creative Destruction”, Federal Reserve Bank of San Francisco Working Paper Series 2017-04 <http://www.frbsf.org/economic-research/publications/working-papers/2017/04/>
- Ahmad, N., J. Ribarsky and M. Reinsdorf (2017), “Can Potential Mismeasurement of the Digital Economy Explain the Post-Crisis Slowdown in GDP and Productivity Growth?”, OECD Statistics Working Papers, 2017/09, OECD Publishing, Paris.
- Ahmad, N. and P. Schreyer (2016), “Measuring GDP in a Digitalised Economy,” OECD Statistics Working Papers, 2016/07, OECD Publishing, Paris.
- Allen, R.G.D. (1949), “The Economic Theory of Index Numbers”, *Economica* 16, 197-203.
- Bean, C. (2016), *Independent Review of UK Economic Statistics*, <https://www.gov.uk/government/publications/independent-review-of-uk-economic-statisticsfinal-report>
- Bils, M. (2009), “Do Higher Prices for New Goods Reflect Quality Growth or Inflation?” *Quarterly Journal of Economics* 124, 637-675.
- Boskin Commission Report (1996), *Toward a More Accurate Measure of the Cost of Living, Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index*. <http://www.ssa.gov/history/reports/boskinrpt.html>
- Brynjolfsson, E., W.E. Diewert, F. Eggers, K.J. Fox and A. Gannamaneni (2018), “The Digital Economy, GDP and Consumer Welfare: Theory and Evidence”, paper presented at the ESCoE Conference on Economic Measurement Bank of England, 16-17 May.
- Brynjolfsson, E., F. Eggers and A. Gannamaneni (2018), “Using Massive Online Choice Experiments to Measure Changes in Well-being”, NBER Working Paper 24514, National Bureau of Economic Research, Cambridge MA, <http://www.nber.org/papers/w24514>.
- Brynjolfsson, E., and J.H. Oh (2012), “The Attention Economy: Measuring the Value of Free Digital Services on the Internet,” Thirty Third International Conference on Information Systems, Orlando 2012.

- Bureau of Labor Statistics (2007), *BLS Handbook of Methods: Ch 17. The Consumer Price Index*. Bureau of Labor Statistics, Washington DC. <https://www.bls.gov/opub/hom/pdf/homch17.pdf>
- Bureau of Labor Statistics (2014), “Quality Adjustment in the Producer Price Index”, U.S. Bureau of Labor Statistics, Washington, D.C., <https://www.bls.gov/ppi/qualityadjustment.pdf>.
- Bureau of Labor Statistics (2018), “Measuring Price Change in the CPI: Telephone hardware, calculators, and other consumer information items”, U.S. Bureau of Labor Statistics, Washington, D.C., <https://www.bls.gov/cpi/factsheets/telephone-hardware.htm>
- Byrne, D. and C. Corrado (2016), “ICT Asset Prices: Marshaling Evidence into New Measures”, Economics Program Working Paper #16-06, The Conference Board, New York, NY.
- Byrne, D. S. Oliner and D. Sichel (2017), “How Fast are Semiconductor Prices Falling?”, *Review of Income and Wealth*, <https://doi.org/10.1111/roiw.12308>
- Byrne, D., J. Fernald and M. Reinsdorf (2016), “Does the United States Have a Productivity Slowdown or a Measurement Problem?” in J. Eberly and J. Stock (eds.), *Brookings Papers on Economic Activity: Spring 2016*, Washington, D.C.: Brookings Institute.
- Diewert, W.E. (1976), “Exact and Superlative Index Numbers”, *Journal of Econometrics* 4, 114-145.
- Diewert, W.E. (1987), “Index Numbers”, pp. 767-780 in J. Eatwell, M. Milgate and P. Newman, (eds.), *The New Palgrave: A Dictionary of Economics*, London: The Macmillan Press.
- Diewert, W.E. (1993a), “The Early History of Price Index Research”, pp. 33-65 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.
- Diewert, W.E. (1993b), “The Economic Theory of Index Numbers: A survey”, pp. 177-228 in *Essays in Index Number Theory*, Volume 1, W.E. Diewert and A.O. Nakamura (eds.), Amsterdam: North-Holland.
- Diewert, W.E. (1998), “Index Number Issues in the Consumer Price Index”, *Journal of Economic Perspectives* 12:1, 47-58.
- Diewert, W.E. and K.J. Fox (2017a), “The Digital Economy, GDP and Consumer Welfare”, paper presented at the CRIW Workshop of the NBER Summer Institute, Cambridge MA, 17-18 July 2017.
- Diewert, W.E. and K.J. Fox (2017b), “Substitution Bias in Multilateral Methods for CPI Construction using Scanner Data,” Vancouver School of Economics Discussion Paper

- 17-02, University of British Columbia.
- Diewert, W.E. and R. Feenstra (2018), “Estimating the Benefits of New Products,” manuscript.
- Feenstra, R.C. (1994), “New Product Varieties and the Measurement of International Prices”, *American Economic Review* 84:1, 157-177.
- Feldstein, M. (2017), “Understanding the Real Growth of GDP, Personal Income, and Productivity”, *Journal of Economic Perspectives* 31, 145-164.
- Goolsbee, A.D. and P.J. Klenow (2018), “Internet Rising, Prices Falling: Measuring Inflation in a World of E-Commerce”, *American Economic Review, Papers and Proceedings* 108, 488-492.
- Groshen, E.L., B.C. Moyer, A.M. Aizcorbe, R. Bradley and D.M. Friedman (2017), “How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies; A View from the Trenches”, *Journal of Economic Perspectives* 31:2, 187-210.
- de Haan, J. (2007), “Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Other Approaches”, Centre for Applied Economic Research Working Paper 2008/01, UNSW.
- de Haan, J. (2017), “Quality Change, Hedonic Regression and Price Index Construction”, forthcoming.
- de Haan, J. and H.A. van der Grient (2011), “Eliminating Chain drift in Price Indexes Based on Scanner Data”, *Journal of Econometrics* 161, 36-46.
- de Haan, J. and F. Krsinich (2012), “The Treatment of Unmatched Items in Rolling Year GEKS Price Indexes: Evidence from New Zealand Scanner Data”, paper presented at the Economic Measurement Group Workshop 2012, Australian School of Business, University of New South Wales, November 23.
- de Haan, J. and F. Krsinich (2014), “Scanner Data and the Treatment of Quality Change in Nonrevisable Price Indexes”, *Journal of Business and Economic Statistics* 32:3, 341-358.
- Haskel, J., P. Goodridge and G. Wallis (2015), “Accounting for the UK Productivity puzzle: A decomposition and predictions,” Discussion Paper 2015/02, Imperial College Business School.
- Hausman, J.A. (1996), “Valuation of New Goods under Perfect and Imperfect Competition”, pp. 20 -236 in *The Economics of New Goods*, T.F. Bresnahan and R.J. Gordon (eds.), Chicago: University of Chicago Press.
- Hausman, J.A. (1999), “Cellular Telephone, New Products and the CPI”, *Journal of Business and Economic Statistics* 17:2, 188-194.

- Hausman, J. (2003), “Sources of Bias and Solutions to Bias in the Consumer Price Index”, *Journal of Economic Perspectives* 17:1, 23-44.
- Hausman, J.A. and G.K. Leonard (2002), “The Competitive Effects of a New Product Introduction: A Case Study”, *Journal of Industrial Economics* 50:3, 237-263.
- Hicks, J.R. (1940), “The Valuation of the Social Income”, *Economica* 7, 105-124.
- Hofsten, E. von (1952), *Price Indexes and Quality Change*, London: George Allen and Unwin.
- Hulten, C. and L. Nakamura (2017), “We See the Digital Revolution Everywhere But in GDP,” presentation to the NBER/CRIW conference on “Measuring and Accounting for Innovation in the 21st Century,” Washington D.C., March 10, 2017.
<http://conference.nber.org/confer/2017/CRIWs17/program.html> (accessed March 10, 2017).
- Ivancic, L., W.E. Diewert and K.J. Fox (2011), “Scanner Data, Time Aggregation and the Construction of Price Indexes”, *Journal of Econometrics* 161, 24-35.
- Keynes, J.M. (1930), *Treatise on Money*, Vol. 1. London: Macmillan.
- Konüs, A.A. (1924), “The Problem of the True Index of the Cost of Living”, translated in *Econometrica* 7, (1939), 10-29.
- Konüs, A.A. and S.S. Byushgens (1926), “K probleme pokupatelnoi cili deneg” (English translation of Russian title: “On the Problem of the Purchasing Power of Money”), *Voprosi Konyunkturi* II(1) (supplement to the Economic Bulletin of the Conjunction Institute), 151-172.
- Marshall, A. (1887), “Remedies for Fluctuations of General Prices’, *Contemporary Review* 51, 355-375.
- Nakamura, L., J. Samuels and R. Soloveichik (2016), “Valuing ‘Free’ Media in GDP: An experimental approach,” paper presented at the Society for Economic Measurement Conference, Thessaloniki, Greece, July 6-8.
- Reinsdorf, M. and P. Schreyer (2017), “Measuring Consumer Inflation in a Digital Economy”, paper presented at the Economic Measurement Workshop 2017, 1 December, UNSW Sydney, Australia.
- Sato, K. (1976), “The Ideal Log-Change Index Number”, *Review of Economics and Statistics* 58, 223-228.
- Syverson, C. (2017), “Challenges to Mismeasurement Explanations for the U.S. Productivity Slowdown”, *Journal of Economic Perspectives* 31, 165-186.
- Tripllett, J. (2004), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*, Directorate for Science, Technology and Industry, DSTI/DOC(2004)9, Paris: OECD.

Vartia, Y.O. (1976), “Ideal Log-Change Index Numbers”, *Scandinavian Journal of Statistics* 3, 121-126.

Wells, J. and A. Restieaux (2014), “Review of Hedonic Quality Adjustment in UK Consumer Price Statistics and Internationally”, U.K. Office for National Statistics, <http://www.ons.gov.uk/ons/guide-method/user-guidance/prices/cpi-and-rpi/review-of-hedonic-quality-adjustment-in-uk-consumer-price-statistics-and-internationally.pdf>