Measuring Inflation under Pandemic Conditions

W. Erwin Diewert
University of British Columbia, UNSW Sydney and NBER

and

Kevin J. Fox
UNSW Sydney

April 2021

Abstract

National statistical offices have faced unprecedented circumstances in the modern history of economic measurement. There were dramatically changing consumer expenditure patterns due to pandemic conditions, with lockdowns and fear of infection making many goods and services unavailable. We examine the implications of changing relative expenditures for the construction of Consumer Price Indexes, with special reference to the treatment of prices for unavailable products. We conclude that for many purposes, it would be useful for statistical agencies to establish a continuous consumer expenditure survey. We also examine various other practical pandemic induced CPI measurement problems.

JEL Classification Numbers: C43, E21, E31

Key Words: Consumer Price Index, disappearing products, pandemic, COVID-19, coronavirus, Cost of Living versus Cost of Goods indexes (COLI versus COGI)

Acknowledgements: The authors thank Paul Armknecht, John Astin, Jonathan Athow, Corinne Becker-Vermeulen, Carsten Boldsen, Kate Burnett-Isaacs, Jan de Haan, François Des Rosiers, David Fenwick, Jason Furman, Daniel Hamermesh, Niall O'Hanlon, Robert Hill, Paul Konijn, Shengyu Li, Jens Mehrhoff, Leigh Merrington, James Poterba, Marshall Reinsdorf, Chihiro Shimizu, Dan Sichel, Mick Silver, Miriam Steurer, Michael Webster, Kam Yu and the referees and editors for helpful comments and discussions on an earlier version of this paper.

W. Erwin Diewert
Department of Economics
University of British Columbia
#997-1873 East Mall
Vancouver, BC V6T 1Z1
CANADA
and NBER
erwin.diewert@ubc.ca

Kevin J. Fox
School of Economics and CAER
UNSW Sydney
NSW 2052
Australia
K.Fox@unsw.edu.au

1. Introduction

The majority of national Consumer Price Indexes (CPIs) are based on pricing out a fixed basket of goods and services that people typically buy in the current period relative to a base period. For National Statistical Offices (NSOs), an implication of the 2020 Covid-19 pandemic and the associated lockdowns was a substantial *disappearing products problem* in the construction of such price indexes. During the pandemic, many goods and services became unavailable and expenditure patterns for still available products changed dramatically; see Carvallo (2020), Carvalho, Garcia, Hansen, Ortiz, Rodrigo, Rodríguez Mora and Ruiz (2020) and Dunn, Hood and Driessen (2020). Thus the fixed basket approach to the construction of a CPI led to measures of consumer price inflation that were likely to be biased because the use of a pre-pandemic basket did not reflect consumer expenditure patterns during the pandemic. This CPI credibility problem was noticed in the financial press. For example:

"Consumption patterns have changed so much that inflation indices are meaningless."
Martin Wolf, *Financial Times*, May 19, 2020.

"But did you notice something about the big price drops quoted? …Great deals are available but no one can take advantage of them. In fact, they're available precisely because no one can take advantage of them…. We have deflation across the basket of goods we usually buy but inflation across the much narrower range of goods we're buying now."
William Watson, *Financial Post*, May 21, 2020.

How to deal with this credibility problem is a major focus of this paper.[1] In addition, we attempt to inform economists who have not specialized in price measurement problems on how actual CPIs are constructed.

[1] We wrote the first version of this paper in April 2020 with the intention of providing statistical agencies with possible methods for dealing with the absence of many goods and services due to pandemic induced lockdowns. Some of our advice has been implemented and some has not.

For products which have disappeared, the advice to NSOs from Eurostat (2020), the International Monetary Fund (2020), the UNECE (2020) and the Intersecretariat Working Group on Price Statistics (2020) [2] was to implement an *inflation adjusted carry forward of missing prices methodology*; i.e., when a price is missing, the price for a commodity for the period prior to the lockdown is used in the current period, with some adjustment for inflation. We show that following this advice can lead to an understatement of inflation relative to a consumer price index that makes use of the concept of a reservation price.

In general, our paper presents a broader review of the price and quantity indexes that statistical agencies could potentially produce during lockdown conditions. The options open to an NSO will depend on its access to current household expenditure data.

The paper is organised as follows. In section 2, we provide a non-technical discussion of different approaches to index number choice and highlight that this paper, for the most part, uses the Fixed Basket approach. The implications of using this approach when the actual consumption basket is changing dramatically under pandemic conditions are discussed. Section 3 looks at comparisons between the Laspeyres, Paasche and Lowe price indexes when inflation adjusted carry forward prices are used for unavailable products, as recommended by international agencies. Section 4 looks at the advantages and disadvantages of using various "practical" price and quantity indexes that statistical agencies are likely to produce during lockdown conditions. We note that the way forward will depend on what types of data are available to the NSO.

Section 5 looks at the problem of a lack of matching product prices at the elementary index level; i.e., we discuss the missing prices problem. Possible methods for dealing with this problem are discussed, depending on the availability of data. Section 6 takes a

[2] The Intersecretariat Working Group on Price Statistics (IWGPS) includes Eurostat, the IMF, ILO, OECD, World Bank and UNECE. It should be noted that the advice that is available on these official websites to deal with the changing weights and missing prices problems is similar to the advice that we suggest in later sections of this paper. The official advice deals with many practical problems that we do not address.

brief look at other practical measurement problems that an NSO may encounter when it attempts to produce a meaningful CPI under pandemic conditions. Section 7 concludes.

2. Alternative Approaches to Index Number Theory

This section presents a brief non-technical overview of the various approaches to index number theory that form the theoretical basis for a CPI. Formal definitions of the relevant indexes follow in the next section.

There are at least four main approaches to index number theory that have been put forth in the literature on bilateral price index numbers. Bilateral index number theory compares prices for two periods. These approaches are as follows:

- *Basket approaches* include the use of Lowe, Laspeyres and Paasche indexes, which will be formally defined in the following section. Informally, a basket index is an index that chooses a representative quantity “basket” and calculates the cost of purchasing this basket at the prices of the current period (numerator) and at the prices of the base period (denominator). The ratio of these two costs forms the price index that calculates inflation going from the base period to the current period. This approach to index number theory is a natural one and goes back hundreds of years. The practical problem with this approach is how exactly are we to choose the “representative” quantity weights? The Lowe index weights prices in each period using the quantities from some base period, which may not be either of the periods being compared. The Lowe index formula is used in CPI construction in most countries.[3] The Laspeyres index uses the quantities from the earlier of the two periods being compared, while the Paasche index using the quantities from the later, or “current”, period. From the viewpoint of index number theory, the Paasche and Laspeyres indexes have the same theoretical footing; i.e., *both are equally representative for the two periods under*

[3] Carsten Boldsen in a personal communication noted that in a 2007 ILO/UNECE survey on the 2004 CPI Manual, of the countries who replied, 65% said that they used a Lowe index while 35% said that they used a Young index. These indexes will be defined in the following section.

consideration. Thus if a single estimate of consumer inflation is to be produced, it is natural to take some sort of average of the Laspeyres and Paasche indexes as a target index for a CPI and it turns out that taking the geometric mean of these two indexes produces a "best" average that will satisfy the time reversal test from the *test approach* to index number choice; see the next dot point below.[4] This geometric mean is the Fisher (1922) index. Thus the basket approach to index number theory (including taking averages of basket type indexes) leads to the Fisher price index as a suitable target index.

- *The Test or Axiomatic approach* to index number theory also leads to the Fisher index. This approach regards the index number formula as a mathematical function of the two price vectors and two quantity vectors that represent the transactions in scope for the two periods under consideration. The approach postulates that the index number formula satisfies various "reasonable" tests or mathematical properties. If a sufficient number of tests are imposed on the index number formula, then an explicit functional form for the index can be determined. An example of an important test is the time reversal test: if the data for the two time periods is exchanged, then the resulting index is the reciprocal of the initial index that did not interchange the data for the two periods. The Fisher and Walsh indexes satisfy this test whereas the Laspeyres and Paasche indexes do not satisfy it. The Fisher index satisfies more reasonable tests than any other commonly used bilateral index number formula (Diewert 1992), and so it is generally regarded as being the "best" index number formula from the viewpoint of the text approach to index number theory.
- *The Stochastic approach* to index number theory. This approach assumes that the price index can be represented as a weighted average of the individual price ratios, say p_n^1/p_n^0 where p_n^t is the price of commodity n in period t for t = 0 (the base period) and t = 1 (the current period). The Törnqvist Theil index is generally regarded as being "best" from this perspective (Theil 1967). This index uses the

[4] Another approach to forming a single representative estimate of inflation is to take the geometric average of the quantity weights for the two periods under consideration as the representative set of quantities. This form of averaging leads to the Walsh (1901) index as a suitable target index. We note that the Walsh and Fisher indexes will generally approximate each other very well; see Diewert (1978; 889).

arithmetic average of expenditure shares in the two periods being compared as the weight for each individual price ratio p_n^1/p_n^0. The overall index is a weighted geometric average of these price ratios using the average shares as weights. It usually approximates the Fisher index closely in empirical applications (Diewert 1978; 889).

- *The Economic Approach or the "Konüs True Cost of Living Index" approach.* This approach brings economics into the picture; i.e., the approach assumes that consumers either maximize utility subject to a budget constraint or they minimize the cost of achieving a certain level of utility. The Fisher, Törnqvist and Walsh indexes all receive an equally strong justification from the perspective of this approach; see Diewert (1976) (1978) (2021a). In the CPI literature, an index based on the economic approach to index number theory is often called a Cost of Living Index or a COLI. It is typically contrasted with a basket type index which is called Cost of Goods Index or COGI; see Deaton and Diewert (2002).

For more on the different approaches to bilateral index number theory, see e.g. ILO (2004), Fisher (1922), Konüs (1924), Theil (1967) and Diewert (1976)(1992)(2021a).

For the most part, we take the Basket Approach to index number theory throughout the paper. We assume that the Laspeyres and Paasche indexes are good basket type indexes: they are intuitively plausible and easy to explain to the public. If they differ substantially and if the situation calls for a single estimate of consumer price inflation, then (from our perspective) these two indexes need to be averaged to give a single credible estimate of inflation, which leads to the Fisher index as their geometric mean. This is regarded as a good index not only from the perspective of the basket approaches to index number theory but it is also a "best" approach from the perspectives of both the economic and test approaches to index number theory.

If quantities consumed do not change much over the two periods being compared, then the actual quantity vectors that are consumed in the two periods being compared may be quite close to the reference quantity vector that is used in the definition of the Lowe index.

There will be some substitution bias (as compared to the economic approach) when using the Lowe index but typically, this "bias" will be relatively small.

However, under pandemic conditions, expenditure patterns changed dramatically. A fixed basket index is very easy to explain and is perfectly reasonable under "normal" conditions. *But a fixed basket index is not particularly relevant when many commodities in the fixed basket are simply not available*. The fixed basket that is implicit in the use of the Lowe index will no longer provide an adequate approximation to actual consumption in period 1. A main message of this paper is that information on actual pandemic expenditure patterns is needed so that "true" Laspeyres, Paasche and Fisher price indexes for the CPI going from pre-pandemic conditions to pandemic conditions can be computed. New estimates of current household expenditures by elementary category are required in order to measure inflation more accurately in the pandemic periods; the old basket weights are almost surely not accurate, even for categories that were not locked down. This point has been demonstrated by recent papers on how consumer expenditures on retail goods have changed due to pandemic conditions; see Cavallo (2020), Carvalho et al. (2020) and Dunn, Hood, and Driessen (2020).

We also encounter an additional significant problem in measuring consumer inflation going from the last pre-lockdown period to the first lockdown period using the Laspeyres or Lowe indexes: *what do we use for prices for the products that are no longer available due to lockdown conditions*? We address this issue in the next section.

3. Alternative Indexes and Inflation Adjusted Carry Forward Prices

Before we address the above question, it will be useful to develop the algebra for the Laspeyres, Paasche and Lowe price indexes between the pre-lockdown period 0 and a subsequent post-lockdown period 1.

We consider CPI goods and services as belonging to one of two groups: Group 1 prices and quantities are available in periods 0 and 1 and Group 2 prices and quantities are only available in period 0.[5] Group 1 products have price and quantity vectors denoted by $p^t \equiv [p_1^t,...,p_M^t] >> 0_M$ and $q^t \equiv [q_1^t,...,q_M^t] >> 0_M$, respectively, for periods t = 0,1.[6] The Group 2 price and quantity vectors for period 0 are $P^0 \equiv [P_1^0,...,P_N^0] >> 0_N$ and $Q^0 \equiv [Q_1^0,...,Q_N^0] >> 0_N$. The Group 2 quantity vector for period 1 is a vector of zero components, so that $Q^1 \equiv 0_N$. The corresponding vector of imputed prices is denoted $P^{1*} \equiv [P_1^{1*},...,P_N^{1*}] >> 0_N$ where $P_n^{1*} > 0$. It is unclear how to define the period 1 price vector P^{1*} for the products that are not available in period 1, where we use the word "product" to cover both goods and services. We will consider the implications of using standard recommended approaches for imputing these prices, but first we will define some index number formulae so it can be seen how these imputed prices come into play in index number construction.

The *Laspeyres* price index going from pre-lockdown period 0 to post-lockdown period 1, P_L^*, is defined as follows:

$$(1)\ P_L^* \equiv [p^1 \cdot q^0 + P^{1*} \cdot Q^0]/[p^0 \cdot q^0 + P^0 \cdot Q^0]$$
$$= s_q^0 P_{Lq} + s_Q^0 P_{LQ}^*$$

where the period 0 expenditure shares for always available commodities and unavailable commodities in period 1 are defined by

$$(2)\ s_q^0 \equiv p^0 \cdot q^0/[p^0 \cdot q^0 + P^0 \cdot Q^0]\ ;\ s_Q^0 \equiv P^0 \cdot Q^0/[p^0 \cdot q^0 + P^0 \cdot Q^0]$$

and the Laspeyres price indexes over always available commodities and unavailable commodities in period 1 are defined by P_{Lq} and P_{LQ} respectively:

$$(3)\ P_{Lq} \equiv p^1 \cdot q^0/p^0 \cdot q^0\ ;\ P_{LQ}^* \equiv P^{1*} \cdot Q^0/P^0 \cdot Q^0.$$

[5] Instead of considering the entire CPI, our discussion can apply to a subset of the CPI. A price index constructed over a subset of the CPI is called an elementary index. If an entire category of consumer expenditures disappears in period 1, this becomes a Group 2 category. The Group 1 category is chosen to be a category that is most closely related to the Group 2 category.

[6] Notation: $p^t >> 0_M$ ($\geq 0_M$) means that all components of the M dimensional vector p^t are positive (nonnegative). The inner product of the vectors p^t and q^t is defined as $p^t \cdot q^t \equiv \Sigma_{m=1}^M p_m^t q_m^t$.

Recall that P^{1*} is a vector of imputed prices for the products that are no longer available in the market place in period 1. We put a superscript asterisk on P^1 to indicate that this price vector is not directly observable. Hence, since the Laspeyres index P_L^* depends on the unobserved vector P^{1*}, we placed a superscript asterisk on the Laspeyres to indicate that these indexes depend on the unobserved price vector P^{1*}. In what follows, we also use this convention for other indexes that similarly depend on P^{1*}.

The *Paasche* and *Lowe* price indexes going from pre-lockdown period 0 to post-lockdown period 1, P_P and P_B^*, are defined as follows:

$$\begin{aligned}(4)\ P_P &\equiv [p^1{\cdot}q^1 + P^{1*}{\cdot}Q^1]/[p^0{\cdot}q^1 + P^0{\cdot}Q^1] \\ &= p^1{\cdot}q^1/p^0{\cdot}q^1 \qquad\qquad \text{since } Q^1 = 0_N; \\ &\equiv P_{Pq}\ ;\end{aligned}$$

$$\begin{aligned}(5)\ P_B^* &\equiv [p^1{\cdot}q^b + P^{1*}{\cdot}Q^b]/[p^0{\cdot}q^b + P^0{\cdot}Q^b] \\ &= s_q^b P_{Bq} + s_Q^b P_{BQ}^*\end{aligned}$$

where the *fixed basket Lowe subindexes* for continuing commodities and unavailable commodities, P_{Bq} and P_{BQ}^* are defined as follows:

$$(6)\ P_{Bq} \equiv p^1{\cdot}q^b/p^0{\cdot}q^b\ ;\ \ P_{BQ}^* \equiv P^{1*}{\cdot}Q^b/P^0{\cdot}Q^b.$$

The *base period hybrid shares* (prices of period 0 but quantities for a prior year b) for the continuing and disappearing commodity groups, s_q^b and s_Q^b, are defined as follows:

$$(7)\ s_q^b \equiv p^0{\cdot}q^b/[p^0{\cdot}q^b + P^0{\cdot}Q^b]\ ;\ s_Q^b \equiv P^0{\cdot}Q^b/[p^0{\cdot}q^b + P^0{\cdot}Q^b].$$

Note that the overall Paasche price index going from the pre-lockdown period 0 to the post-lockdown period 1, P_P, defined by the first equation in (4) turns out to equal the Paasche subindex, P_{Pq}, that uses only the prices for the commodities that are available in periods 0 and 1. Thus the "true" overall Paasche price index is equal to the Paasche price index for always available commodities (and thus both indexes are observable in principle).

The situation is different for the overall Laspeyres and Lowe price indexes, P_L^* defined by (1) and P_B^* defined by (5); *these indexes require estimates for the product prices that are missing in period 1*; i.e., we require an estimate for the vector of period 1 prices, P^{1*}, in order to calculate these indexes.

What is the "right" price for missing product n in period 1, P_n^{1*}? At first glance, one could argue that the "right" price is $P_n^{1*} \equiv +\infty$; i.e., it is impossible to purchase product n in period 1 at any finite price so a market price that will ensure that no one will purchase product n in period 1 is a price that is infinitely high. However, this is where economic theory can play a role. Normally, market prices are determined by the intersection of supply and demand curves. In any given period, the intersection of these curves determines P_n^1 and Q_n^1. But lockdowns of some parts of the economy mean that some products n simply become unavailable in period 1. What has happened is that the *supply curve* for product n has become straight line that is parallel to the price axis and this line has shifted to become identical to the vertical price axis.[7] Thus the (unobserved) market price for the product n under consideration is the price where the demand curve intersects the vertical price axis; this determines P_n^{1*} conceptually. This price will typically be less than $+\infty$; i.e., it does not take an infinite price to deter households from purchasing any particular product. This intersection price P_n^{1*} can be interpreted as a *Hicksian reservation price* in a Cost of Living Index (COLI) context. Hicks (1940; 114) introduced

[7] In addition to government mandated lockdowns, the unavailability of a commodity may be due to a lack of demand which causes supplying firms to shut down. This is where the difference between a COLI and a COGI comes into play. If the lack of supply for a commodity is due to a shift in consumer preferences (due to safety concerns), then from a COLI point of view, we need to compute two (conditional) Cost of Living Indexes: one that uses the preferences of the pre-pandemic period and one that uses the preferences of the post pandemic period. From a COGI point of view, changing preferences are irrelevant: what matters is the intersection of aggregate demand and aggregate supply curves to determine the market prices and quantities which should be used in a COGI. Thus we interpret the appropriate COGI price for an unavailable product to be the price where the relevant market demand curve intersects the price axis. However the determination of this unobserved "market" price necessarily involves some econometric modeling and hence we cannot expect statistical agencies to be able to produce estimates for these prices in real time.

the concept of a reservation price into the economics literature and Hofsten (1952) introduced the term. Reservation prices have been widely used by economists to measure the benefits of new products since the pioneering work of Feenstra (1994). Reservation prices can be measured retrospectively; see Hausman (1996) (1999) (2003) and Diewert and Feenstra (2019) for empirical examples.

Below, we will refer to an index which uses reservation prices for the missing prices as a "true" index. In the literature on the economic approach to index number theory, it is common to refer to a COLI as the "true" index. This convention dates back to a paper written by Robert Pollak for the US Bureau of Labor Statistics in 1971, which was reprinted as Pollak (1983).

Currently, it is unlikely that national statistical offices can estimate such reservation prices, at least not in a timely fashion using presently available techniques. We outline below how the algebra for the Laspeyres and Lowe indexes works if inflation adjusted carry forward prices are used instead.

Define the *inflation adjusted carry forward prices* for the missing products in period 1, P^{1L}, using the Laspeyres index for always present products, P_{Lq}, as the inflation index[8] as follows:

$$(8)\ P^{1L} \equiv P_{Lq}P^0 = (p^1{\cdot}q^0/p^0{\cdot}q^0)P^0.$$

These imputed prices can be used as replacement prices for the reservation prices P^{1*} in definition (1) for the overall Laspeyres price index. We called the resulting index P_L^{CL}, which is a Laspeyres inflation adjusted carry forward index using the Laspeyres index P_{Lq} as the adjusting inflation index. Thus we have:

[8] National Statistical Agencies do not in general use P_{Lq} as the carry forward inflation index; they use a wide variety of alternative indexes such as a single price ratio of a close substitute product for the missing product. See Eurostat (2020), IMF (2020), UNECE (2020) and IWGPS (2020) for lists of possible alternative indexes.

$$
\begin{aligned}
(9)\ P_L^{CL} &\equiv [p^1 \cdot q^0 + P^{1L} \cdot Q^0]/[p^0 \cdot q^0 + P^0 \cdot Q^0] \\
&= [p^1 \cdot q^0 + (p^1 \cdot q^0/p^0 \cdot q^0)P^0 \cdot Q^0]/[p^0 \cdot q^0 + P^0 \cdot Q^0] && \text{using definition (8)} \\
&= P_{Lq} \\
&= P_{Lq}(s_q^0 + s_Q^0) && \text{since } s_q^0 + s_Q^0 = 1 \\
&= s_q^0 P_{Lq} + s_Q^0 P_{Lq} \\
&< s_q^0 P_{Lq} + s_Q^0 P_{LQ}^* && \text{if } P_{LQ}^* > P_{Lq} \\
&= P_L^* && \text{using definition (1).}
\end{aligned}
$$

The above equations tell us two things:

- The overall Laspeyres index that is generated by the use of the carry forward prices P^{1L} defined by (8) gives rise to an index P_L^{CL} which turns out to be exactly equal to P_{Lq}, the Laspeyres price index that is restricted to always available products;
- The inflation adjusted carry forward Laspeyres price index, P_L^{CL} defined by (9), will be less than the "true" overall Laspeyres price index P_L^* provided that the "true" Laspeyres index defined over the Group 2 products, P_{LQ}^*, is greater than the Laspeyres index defined over always available products, P_{Lq}.

A sufficient condition for $P_{LQ}^* > P_{Lq}$ is the following one:

$$(10)\ P_n^{1*} > P_{Lq}P_n^0\,; \qquad n = 1,...,N.$$

Conditions (10) are that the period 1 reservation prices for unavailable products P_n^{1*} are greater than the corresponding inflation adjusted prices from period 1, $P_{Lq}P_n^0$, for each unavailable product n = 1,...,N. This is likely to be the case.

However, it is possible that $P_n^{1*} = P_{Lq}P_n^0$ for some products n. In a simplified scenario where we have only one unavailable product and one always available product, this case occurs if the two products are *perfect substitutes*; i.e., the consumer's utility function is linear in the two products. Put another way, in terms of our simple demand equals supply partial equilibrium approach to the determination of reservation prices, this perfect substitutes case corresponds to the case where the demand curve for product n is parallel to the Q_n axis. Thus as the supply curve shifts to a vertical straight line that coincides with the P_n axis, the price of the product remains constant after adjustment for general inflation in the always available goods and services. It is unlikely that this flat demand curve scenario could approximate actual consumer behavior during a pandemic; it is much more likely that demand curves are downward sloping and in this case, we will get a downward bias in the inflation adjusted carry forward Laspeyres index relative to the "true" Laspeyres index. As indicated earlier, it is common in the COLI literature to call a COLI the "true" index. But a Laspeyres index is not a COLI except under strong assumptions (of no substitution between products). Our "true" Laspeyres index is simply a Laspeyres index that uses our concept of (unobserved) market clearing prices for unavailable products. Similarly for the "true" Lowe index.

The algebra for the Lowe index is much the same as the above algebra for the true Laspeyres index. The "true" Lowe index was defined by (5)-(7) above. The inflation adjusted carry forward prices P^{1I} for the missing period 1 products are now defined as follows:

(11) $P^{1I} \equiv P_{Bq}P^0 = (p^1 \cdot q^b / p^0 \cdot q^b)P^0.$

Substitute the inflation adjusted carry forward prices P^{1I} defined by (11) into definition (5) in order to obtain the following approximation, P_B^{CI}, to the "true" Lowe index, P_B^*:

(12) $P_B^{CI} \equiv [p^1 \cdot q^b + P^{1I} \cdot Q^b]/[p^0 \cdot q^b + P^0 \cdot Q^b]$

$$\begin{aligned}
&= [p^1 \cdot q^b + (p^1 \cdot q^b/p^0 \cdot q^b)P^0 \cdot Q^b]/[p^0 \cdot q^b + P^0 \cdot Q^b] && \text{using definition (8)} \\
&= P_{Bq} \\
&= P_{Bq}(s_q^0 + s_Q^0) && \text{since } s_q^0 + s_Q^0 = 1 \\
&= s_q^0 P_{Bq} + s_Q^0 P_{Bq} \\
&< s_q^0 P_{Bq} + s_Q^0 P_{BQ}^* && \text{if } P_{BQ}^* > P_{Bq} \\
&= P_B^* && \text{using definition (5).}
\end{aligned}$$

The above equations have the following implications:

- The overall Lowe index that is generated by the use of the carry forward prices P^{1I} defined by (11) gives rise to an index P_B^{CI} which turns out to be exactly equal to P_{Bq}, the Lowe price index that is restricted to always available products;
- The inflation adjusted carry forward Lowe price index, P_B^{CI} defined by (12), will be less than the "true" overall Lowe price index P_B^* provided that the "true" Lowe index defined over the Group 2 products, P_{BQ}^*, is greater than the Lowe index defined over always available products, P_{Bq}.

A sufficient condition for $P_{BQ}^* > P_{Bq}$ is the following one:

$$(13)\ P_n^{1*} > P_{Bq}P_n^0\ ; \qquad n = 1,...,N.$$

As before, we think that it is extremely likely that inflation adjusted carry forward prices $P_{Bq}P_n^0$ are well below the corresponding market clearing reservation prices P_n^{1*} and thus there is a very high probability that the inflation adjusted carry forward Lowe index P_B^{CI} defined by (12) understates our suggested definition of the "true" Lowe index P_B^* defined by (5).

We acknowledge that statistical offices will not be in a position to calculate satisfactory approximations to the "true" Lowe index, P_B^*, which measures inflation going from a pre-lockdown period to a post-lockdown period. In which case, it seems reasonable that NSOs notify users of their data that the price index that they put out in the first lockdown period (some version of the Lowe index that uses some version of carry forward prices) is unlikely to be an accurate measure of inflation that is comparable to previous index values.[9] The comparison will be somewhat accurate for the subindex that is restricted to

[9] This suggestion is somewhat moot at this time of publication. However, we wrote the first version of this paper in April 2020 when it was not clear what NSOs were going to do with their CPIs.

always available products. The problem is the fact that the Lowe pre-lockdown basket weights, (q^b, Q^b), may not be close to the post-lockdown weights, $(q^1,0_N)$.

As indicated in the previous section, a Lowe index will approximate a Laspeyres index and if the quantity weights do not change much going from period to period, so that $q^b \approx q^0 \approx q^1$, then the Lowe subindex for continuing products, P_{Bq}, will provide an adequate approximation to the Laspeyres and Paasche subindexes, P_{Lq} and P_{Pq}. However, empirical evidence suggests that expenditure shares on food and other available products did not remain approximately constant in the first lockdown period so that q^1 was substantially different from q^0. Hence, once information on current period expenditures becomes available, it would be useful to compute Fisher subindexes over always available products on a retrospective basis for the pandemic affected periods. The retrospective index could be called an analytic index or a supplementary index. We suggest that the retrospective index be a Fisher index if possible (for levels of aggregation where current period price and expenditure information is available) because of the good test properties of the Fisher formula.

The above algebra applies to indexes that are calculated using prices and quantities (or equivalently, using expenditures and unit value prices). At higher levels of aggregation, the prices become elementary price indexes for commodity categories and the quantities become *volumes*; i.e., they are expenditures deflated by the relevant price indexes. The above algebra applies in both situations.

Many countries use a Young index instead of the Lowe index as a target index in the production of their CPIs at higher levels of aggregation. Recall that the Lowe index made use of the base period quantity vectors $q^b \equiv [q_1^b,\ldots q_M^b]$ and $Q^b \equiv [Q_1^b,\ldots Q_N^b]$. In order to define the Young index, we need to make use of the base period expenditure share vectors, $s^b \equiv [s_1^b,\ldots s_M^b]$ and $S^b \equiv [S_1^b,\ldots S_N^b]$ where $\sum_{m=1}^M s_m^b + \sum_{n=1}^N S_n^b = 1$. Using the above notation for observed prices p_m^t and P_n^t in each period t, the "true" Young index is defined as follows:

(14) $P_{Yb}{}^* \equiv \sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0) + \sum_{n=1}^{N} S_n{}^b(P_n{}^{1*}/P_n{}^0)$.

It can be seen that the Young index is a member of the class of stochastic or descriptive statistics indexes. If the base period b happens to be period 0, then it can be seen that $P_{Yb}{}^*$ becomes $P_{Y0}{}^* = P_L{}^*$, the "true" Laspeyres index defined earlier by (1). This choice for the base period b share weights leads to a *relevant* Young index. Another *relevant* choice for the period b expenditure shares are the expenditure shares of period 1. The resulting Young index is equal to P_{Y1} defined as follows:

(15) $P_{Y1} \equiv \sum_{m=1}^{M} s_m{}^1(p_m{}^1/p_m{}^0) + \sum_{n=1}^{N} S_n{}^1(P_n{}^{1*}/P_n{}^0) = \sum_{m=1}^{M} s_m{}^1(p_m{}^1/p_m{}^0)$

where the second equality in (15) follows since the expenditure shares for the unavailable products $S_n{}^1$ are all equal to 0. The Young index P_{Y1} defined by (15) is closely related to the overall Paasche index, P_P, and the Paasche index restricted to always available products, P_{Pq}, defined earlier by (4); i.e., we have the following equalities and inequality:

(16) $P_P = P_{Pq} = [\Sigma_{m=1}^{M} s_m{}^1(p_m{}^1/p_m{}^0)^{-1}]^{-1} \leq \sum_{m=1}^{M} s_m{}^1(p_m{}^1/p_m{}^0) = P_{Y1}$

where the inequality follows since a weighted harmonic mean is always equal to or less than the corresponding weighted arithmetic mean; see Hardy, Littlewood and Polya (1934; 26). Typically, the gap between P_P and P_{Y1} will not be large.

$P_{Y0}{}^*$ and P_{Y1} are equally relevant Young indexes which could be used to measure the price change that occurred between periods 0 and 1. Thus some form of averaging of these two indexes is called for if a single measure of inflation is desired. The geometric mean of $P_{Y0}{}^*$ and P_{Y1} is a useful average that leads to an index which satisfies the important time reversal test. However, statistical agencies will not be able to calculate $P_{Y0}{}^*$ in real time due to the difficulty in determining the market clearing unobserved prices P^{1*}. Thus to form a real time Young index, it will be necessary for statistical agencies to use some form of inflation adjusted carry forward prices.

Recall that the vector of inflation adjusted carry forward prices for the Laspeyres index, P^{1L}, was defined above by (8): $P^{1L} \equiv P_{Lq}P^0 = (p^1\cdot q^0/p^1\cdot q^0)P^0$. An analogous vector of *inflation adjusted carry forward prices for the missing products* in period 1 for the Young index $P_{Yb}{}^*$ is P^{1Y} defined as follows:

$$(17)\ P^{1Y} \equiv \{[\sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0)]/[\sum_{m=1}^{M} s_m{}^b]\}P^0.$$

Note that the inflation index used to adjust the base period prices P^0 into imputed prices for missing commodities in period 1 is $[\sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0)]/[\sum_{m=1}^{M} s_m{}^b]$ which is essentially a Young index for the always available commodities where the base period shares for always available products $s_m{}^b$ have been reweighted so that they sum to one. Now substitute the imputed carry forward prices P^{1Y} defined by (17) into definition (14) in order to obtain an *inflation adjusted carry forward Young index* $P_{Yb}{}^{CF}$:

$$\begin{aligned}(18)\ P_{Yb}{}^{CF} &\equiv \sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0) + \sum_{n=1}^{N} S_n{}^b(P_n{}^{1Y}/P_n{}^0)\\
&= \sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0) + \{[\sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0)]/[\sum_{m=1}^{M} s_m{}^b]\}\{\sum_{n=1}^{N} S_n{}^b]\\
&= [\sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0)]\{1 + [\sum_{n=1}^{N} S_n{}^b]/[\sum_{m=1}^{M} s_m{}^b]\}\\
&= [\sum_{m=1}^{M} s_m{}^b]^{-1}[\sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0)][\sum_{m=1}^{M} s_m{}^b + \sum_{n=1}^{N} S_n{}^b]\\
&= \sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0)/\sum_{m=1}^{M} s_m{}^b.\end{aligned}$$

Thus using the inflation adjusted carry forward prices defined by (17) for the missing prices will cause the Young index $P_{Yb}{}^{CF}$ to collapse down to the reweighted Young index that measures price change for just the always available products. If the "true" market clearing period 1 prices, $P_n{}^{1*}$ are greater than the inflation adjusted prices $P_n{}^{1Y} \equiv \{[\sum_{m=1}^{M} s_m{}^b(p_m{}^1/p_m{}^0)]/[\sum_{m=1}^{M} s_m{}^b]\}P_n{}^0$ for $n = 1,\ldots,N$, then the "true" Young index will be greater than the inflation adjusted Young index defined by (18).

4. The Way Forward

As we have noted, statistical agencies that use a fixed basket methodology for constructing their CPI are faced with the fact that the fixed basket is no longer as relevant for pandemic periods as it was in pre-pandemic times. Thus as we have indicated in the previous section, the use of Lowe or Young indexes in pandemic times that use pre-pandemic weights will not accurately reflect changes in the cost of purchasing goods and services during pandemic periods. However, many NSOs will not have the resources to estimate representative baskets in real time. We will list a number of possible strategies that an agency could use in order to construct a CPI under pandemic conditions, depending on what kind of data they are able to collect. We will start with the assumption that very little data are available and finish with the way forward if ample data are available. For each of these cases, we will look at possible ways of addressing the lack of matching problem at the elementary index level.

Case 1: Very Little Data Availability

For this case, we suppose that the agency has only a fixed basket (q^b,Q^b) along with price data for period 0 which is the period before the lockdown, (p^0,P^0). For the pandemic periods, the agency has only price data for always available goods and services, p^t for $t = 1,2,...\tau-1$, the pandemic periods. When the pandemic is over in period τ, we assume that the agency can collect price data for always available goods, p^τ, and for commodities that were available in period 0 and become available again in period τ, P^τ. For the lockdown periods, the agency can calculate the fixed basket index for always available commodities, $p^t\cdot q^b/p^0\cdot q^b$ for $t =1,2,...\tau-1$. These indexes may be suitable for (partial) compensation purposes; i.e., if period 0 household expenditures on the basket q^b were equal to $p^0\cdot q^b$, then using the index $p^t\cdot q^b/p^0\cdot q^b$ to escalate the household's period 0 "income" (equal to $p^0\cdot q^b$) would allow the household to purchase the bundle of commodities q^b in period t for $t = 1,2,...,\tau-1$. This index may be subject to some substitution bias. From the COLI perspective, the NSO would need to note that the CPI for these periods is not comparable to the CPI for either period 0 or period τ.[10] To provide a useful estimate for a cost of living index relative to the standard of living in period 0 for the lockdown periods, we require estimates for reservation prices, P^{t*} for $t = 1,2,...,\tau-1$. Very few NSOs will venture to estimate reservation prices. What NSOs can do is to provide a credible CPI for goods and services which are actually available during the lockdown periods. When the lockdown ends and conditions approach "normality" in period τ, then the under-resourced statistical office can use its pre-lockdown basket, (q^b,Q^b), to calculate the price level in period τ relative to period 0 as the fixed base index $[p^\tau \cdot q^b + P^\tau \cdot Q^b]/[p^0 \cdot q^b + P^0 \cdot Q^b]$.

Case 2: Some Data Availability

We assume that the data availability is at least as good as in the above case. In addition, we assume that by period σ (where $1 < \sigma < \tau$), the statistical agency is able to obtain an estimate for a representative quantity vector q^σ for the always available quantities during the lockdown period. For the lockdown periods prior to period σ, the agency can calculate the fixed basket index for always available commodities, $\pi^t \equiv p^t \cdot q^b/p^0 \cdot q^b$ for $t = 1,2,...\sigma-1$. In period σ, the new basket q^σ becomes available so it is possible to calculate the period t price index value for period t as $\pi^t \equiv \pi^{t-1}[p^t \cdot q^\sigma/p^{t-1} \cdot q^\sigma]$ for $t = \sigma,\sigma+1,...,\tau-1$. However, the price levels π^1, π^2,..., $\pi^{\sigma-1}$ may very unreliable due to the fact that the pre-lockdown quantity vector q^{b0} may be rather far from the actual consumption vectors q^1, q^2,...,$q^{\sigma-1}$ over the lockdown period extending from period 1 to period $\sigma-1$. Thus it may be preferable to define ρ^σ as the pseudo Fisher index comparing period 0 with period σ; i.e., define $\pi^\sigma \equiv [p^\sigma \cdot q^b/p^0 \cdot q^b]^{1/2}[p^\sigma \cdot q^\sigma/p^0 \cdot q^\sigma]^{1/2}$. For lockdown periods following period σ but prior to period τ, define $\pi^t \equiv \pi^{t-1}[p^t \cdot q^\sigma/p^{t-1} \cdot q^\sigma]$ for $t = \sigma+1,\sigma+2,...,\tau-1$. Again, these indexes may be suitable for partial indexation purposes but they are likely to substantially understate "true" COLI-type inflation relative to the pre-pandemic period. When we get

[10] Equations (12) show that the indexes $p^t \cdot q^b/p^0 \cdot q^b$ for $t = 1,2,...\tau-1$ could be generated by using the inflation adjusted carry forward price vectors for unavailable commodities defined as $P^{tI} \equiv (p^t \cdot q^b/p^0 \cdot q^b)P^0$. If this is done, users need to be informed that the resulting indexes are not "true" fixed basket indexes in that part of the overall fixed basket, (q^b,Q^b), is simply not available for purchase in period t. A similar comment applies to NSOs using Young indexes.

to period τ, the moderately-resourced statistical office can calculate the fixed base index relative to period 0; i.e., set $\pi^\tau = [p^\tau \cdot q^b + P^\tau \cdot Q^b]/[p^0 \cdot q^b + P^0 \cdot Q^b]$.

If the statistical office has set in motion a continuous consumer expenditure survey so that a new period τ comprehensive basket (q^τ,Q^τ) can be constructed, then the office can calculate the pseudo Fisher indexes defined above. If the office has access to scanner data for some strata, then Fisher indexes can be calculated for those strata.

Case 3: Ample Data Availability

We assume the Case 1 data availability plus the availability of representative quantity vectors q^{bt} for all periods $t = 0,1,...,\tau$. We also assume that a representative quantity vector for the unavailable commodities is available for periods 0 and τ. Denote these vectors by Q^{b0} and $Q^{b\tau}$. The corresponding price vectors are P^0 and P^1. For period 0, define the price level as $\pi^0 \equiv 1$. For the lockdown periods, define the period 1 price index π^1 as the pseudo Fisher index $\pi^1 \equiv \{[p^1 \cdot q^{b0}/p^0 \cdot q^{b0}][p^1 \cdot q^{b1}/p^0 \cdot q^{b1}]\}^{1/2}$. For $t = 2,3,...\tau-1$ define the period t price index as $\pi^t \equiv \pi^{t-1}\{[p^t \cdot q^{bt-1}/p^{t-1} \cdot q^{bt-1}][p^t \cdot q^{bt}/p^{t-1} \cdot q^{bt}]\}^{1/2}$. Thus the period to period pseudo Fisher indexes are chained together to form the period t price level. For period τ, define the price level π^τ as the comprehensive pseudo Fisher price index connecting period 0 to period τ; i.e., define π^τ as follows:

$$(19)\ \pi^\tau \equiv \{[p^\tau \cdot q^{b0} + P^\tau \cdot Q^{b0}]/[p^0 \cdot q^{b0} + P^0 \cdot Q^{b0}]\}^{1/2}\{[p^\tau \cdot q^{b\tau} + P^\tau \cdot Q^{b\tau}]/[p^0 \cdot q^{b\tau} + P^0 \cdot Q^{b\tau}]\}^{1/2}.$$

The reason for using chained pseudo Fisher price indexes for the available products during the lockdown period instead of fixed base pseudo Fisher price indexes is the likelihood that consumer purchases of available products over the lockdown periods may not be well approximated by a constant vector q^b. Initially, households will stock up on storable goods and cut back on purchases of consumer durables. If the lockdown period is long and the degree of lockdown varies, then it is quite likely that the vector of actual purchases of available commodities in period t, q^t, will be quite variable and hence a constant q^b will not provide a representative vector of household purchases over all of the lockdown periods. Note that the set of available products has varied over lockdown periods. In the early stages, food manufacturers did not produce their full line of products; they concentrated on increasing the volume of their best selling products and produced them at scale to satisfy stockpiling demands. Of course, the gold standard for the quantity vectors q^{bt} would be the actual period t consumption vectors, q^t, in which case, the pseudo Fisher indexes would become actual Fisher indexes.

5. The Lack of Matching Problem at the Elementary Index Level

A problem which has appeared as a result of country wide lockdowns is the *problem of missing products and services* in retail outlets. In many cases, the missing products and services reappeared in a later period; in some cases, they were gone for the duration of the lockdown. If the products are gone for the duration of the lockdown and the remaining products are present during the current and prior lockdown periods, then we are in position to apply the theory above to the particular elementary aggregate under consideration; i.e., we need to switch from pricing out the pre-lockout basket of products to the new smaller set of products. However, real life will be more complicated than having a clear division between products present and products that have been discontinued for all lockout periods: products will be drifting in and out of scope in any particular retail outlet. This may lead to a massive lack of matching problem. We will briefly suggest possible solutions to this problem under two scenarios: (i) only web scraped data are available and (ii) scanner data are available. The analysis in this section differs from the analysis that was presented in the previous section where it was known that some commodities would be unavailable for the duration of the lockdown. We now assume that the full array of pre-lockdown products is not available in the lockdown periods.

Case 1: Only Price Data are Available

Method 1: Adapt the Section 3 Carry Forward Methodology

The adaptation here is to assume that q^0 and q^1 are equal to the vector of ones, 1_M, and Q^0 equals the vector of ones, 1_N. $Q^1 = 0_N$ as in section 4. Thus the q group of products are the *maximum overlap products* that are present in both periods and the Q products are present in the base period 0 but not in the current period 1. The given price vectors are p^0, p^1 and P^0. Applying the section 4 methodology using the above assumptions on prices and quantities leads to the following inflation adjusted carry forward price index using equation (12) adapted to the present situation:

$$(20)\ P_B^{CI} = P_{Bq} = p^1 \cdot 1_M / p^0 \cdot 1_M = \Sigma_{m=1}^{M} p_m^{1} / \Sigma_{m=1}^{M} p_m^{0}.$$

The above index is the Dutot (1738) elementary index, defined over products that are present in both periods. It has an undesirable property: it is not invariant to changes in the units of measurement of the products. It will also give a higher weight to products that are more expensive which may not be a desirable property. Nevertheless, it does approximate the theoretically more desirable Jevons index under certain conditions (Diewert 2021b).

It is a standard practice to use a sample of prices to represent price movements for the universe of commodities in an elementary category of transactions. This can be problematic during a pandemic. For example, during the pandemic, international travel by airlines fell dramatically but some flights still took place. Thus statistical agencies could use the price movements for the existing flights to represent the movement of prices over the entire air travel universe. But this practice disguises the fact that international air travel for most households was shut down for long periods of time. Thus the air travel category should be split into at least two categories in a CPI: one category for some households who are allowed to travel internationally and another category for the locked down households. The types of "bias" that we discussed in section 3 are applicable to this situation. This potential bias problem is present whenever a sample of prices is used to represent movements in the universe of prices in scope. However, the "bias" will usually be small (from the viewpoint of the economic approach to index number theory) if the commodities in the category are highly substitutable with each other. In this case, consumers can easily substitute towards the available varieties when some varieties disappear without much overall loss of utility (Diewert 2021d). The pandemic situation is very different: the pandemic induced disappearance of many commodities surely led to large losses in utility, which were not measured. There is little that statistical offices can do to remedy this situation but it seems reasonable for NSOs to flag this problem.

Method 2: Use Maximum Overlap Jevons Indexes

This method simply sets the price index equal to the Jevons (1865) index for the overlapping products in the two periods under consideration. Thus using the same notation as was used to describe Method 1 above, the maximum overlap Jevons index, P_{JMO}, is equal to the geometric mean of the price ratios for the overlapping products:

(21) $P_{JMO} \equiv [\Pi_{m=1}{}^{M} (p_m{}^1/p_m{}^0)]^{1/M}$.

The Jevons index has the best axiomatic properties for indexes (with no missing prices) that depend only on prices. Note in particular that the maximum overlap Jevons index is invariant to changes in the units of measurement for the products (Diewert 2021b).

Method 3: Use the Multilateral Time Product Dummy Method

A problem with the above two methods is that they make use of price data covering only two periods. In the situation where closely related products are moving in and out of scope, constructing maximum overlap bilateral index numbers does not make use of all of the data and hence is inefficient from a statistical point of view. For example, suppose a product is present in periods 1 and 3 and another product is present in periods 2 and 4. In a bilateral index setup, the information pertaining to these two products would not be used which is inefficient since price comparisons for product 1 between periods 1 and 3 and for product 2 between periods 2 and 4 are perfectly valid comparisons and should be used somehow in constructing the sequence of price indexes. The way forward here is to use a *multilateral index* which utilizes the price information for all periods. For studies on the use of multilateral indexes in the time series context, see Balk (1980), Ivancic, Diewert and Fox (2011) and de Haan and van der Grient (2011). For a detailed discussion on the use of multilateral indexes in the time series context, see Diewert (2021c).

A widely used multilateral method is the *Time Product Dummy Method* with missing observations. The method was originally devised for making price comparisons across countries and is known as the Country Product Dummy multilateral method; see Summers (1973). A weighted version of this model (with missing observations) was first applied in the time series context by Aizcorbe, Corrado and Doms (2000). The method can be interpreted as a special case of a hedonic regression model; see de Haan (2004) (2010) and de Haan and Krsinich (2014) (2018).

We introduce some new notation in order to describe this method. We now assume that there are N products and T time periods but not all products are purchased (or sold) in all time periods. The price and quantity vectors for period t are denoted by $p^t \equiv [p_{t1},...,p_{tN}]$ and $q^t \equiv [q_{t1},...,q_{tN}]$. If product n in period t is missing, we set the corresponding price and quantity, p_{tn} and q_{tn}, equal to 0. For each period t, define the set of products n that are present in period t as $S(t) \equiv \{n: p_{tn} > 0\}$ for t = 1,2,...,T. It is assumed that these sets are not empty; i.e., at least one product is purchased in each period. For each product n, define the set of periods t where product n is present as $S^*(n) \equiv \{t: p_{tn} > 0\}$. Again, assume that these sets are not empty; i.e., each product is sold in at least one time period. Define the integers N(t) and T(n) as follows:

$$(22)\quad N(t) \equiv \Sigma_{n\in S(t)}\, 1; \qquad t = 1,...,T;$$

$$(23)\quad T(n) \equiv \Sigma_{t\in S^*(n)}\, 1; \qquad n = 1,...,N.$$

If all N products are present in period t, then N(t) = N; if product n is present in all T periods, then T(n) = T.

The economic model that is consistent with the Time Dummy Product multilateral method is the following one:

(24) $p_{tn} = \pi_t\alpha_n$; $\qquad t = 1,...,T;\ n\in S(t)$

where π_t is the period t price level and α_n is a quality adjustment parameter for product n. If all products were available in all periods, equations (24) tell us that prices for the group of products in scope are moving in a proportional manner. This is consistent with purchasers of the N products having the linear utility function, $f(q) = \alpha\cdot q \equiv \Sigma_{n=1}^{N} \alpha_n q_n$ where $\alpha \equiv [\alpha_1,...,\alpha_N]$ and $q \equiv [q_1,...,q_N]$. It can be seen that this approach will only be adequate if the products are very close substitutes since a linear utility function implies that the products are perfect substitutes; see Diewert (2021d) for further explanation of the underlying economic model.

Now take logarithms of both sides of equations (24), add error terms e_{tn} to the resulting equations and we obtain the following system of estimating equations:

(25) $\ln p_{tn} = \rho_t + \beta_n + e_{tn}$; $\qquad t = 1,...,T;\ n\in S(t)$

where $\rho_t \equiv \ln\pi_t$ for $t = 1,...,T$ and $\beta_n \equiv \ln\alpha_n$ for $n = 1,...,N$. Note that equations (25) form the basis for the *time dummy hedonic regression model*. This is Court's (1939; 109-111) hedonic suggestion number two. He chose to transform equations (24) by the log transformation because the resulting regression model fit his data on automobiles better. Diewert (2003) also recommended the log transformation on the grounds that multiplicative errors were more plausible than additive errors.

Estimates for the unknown parameters ρ_t and β_n that appear in equations (25) can be found by solving the following least squares minimization problem:

(26) $\min_{\rho,\beta} \{\Sigma_{t=1}^{T} \Sigma_{n\in S(t)} [\ln p_{tn} - \rho_t - \beta_n]^2\}$.

In order to obtain a unique solution to (26), we need to impose a full rank condition on the X matrix generated by the linear regression model defined by equations (25) and ρ_1 = 0 (Diewert 2021c), and impose a normalization on the parameters. Choose the normalization $\rho_1 = 0$ (which corresponds to $\pi_1 = 1$). Denote the resulting solution by $\rho^* \equiv [1,\rho_2^*,...,\rho_T^*]$ and $\beta^* \equiv [\beta_1^*,...,\beta_N^*]$. Use these estimates to form estimates for $\pi_t^* \equiv \exp[\rho_t^*]$ for t = 1,...,T and $\alpha_n^* \equiv \exp[\beta_n^*]$ for n = 1,...,N. It turns out that these estimates satisfy the following equations:

(27) $\pi_t^* = \Pi_{n\in S(t)} [p_{tn}/\alpha_n^*]^{1/N(t)}$; t = 1,...,T;

(28) $\alpha_n^* = \Pi_{t\in S^*(n)} [p_{tn}/\pi_t^*]^{1/T(n)}$; n = 1,...,N.

Note that p_{tn}/α_n^* is a *quality adjusted price* for product n in period t and p_{tn}/π_t^* is the corresponding *inflation adjusted price* for product n in period t. Thus the period t estimated price level, π_t^*, is the geometric mean of the quality adjusted prices for products that are available in period t and the estimated quality adjustment factor for product n, α_n^*, is the geometric mean of all of the inflation adjusted prices for product n over all periods. Note that if the set of available products in periods r and t is the same, then $\pi_t^*/\pi_r^* = [\Pi_{n\in S(t)}(p_{tn}/p_{rn})]^{1/N(t)}$ which is the Jevons index defined over the products that are present in both periods. These price levels generated by this method have satisfactory axiomatic properties; see Diewert (2021c). It turns out that the price levels satisfy an *identity test* so if prices are equal in periods r and t, then $\pi_r^* = \pi_t^*$. There are some additional choices that the statistical agency will have to make if it uses this method; i.e., it is necessary to decide on the length of the window of observations T and it is necessary to decide on how to link the results of the latest window of estimates with the previous window of estimates for the price levels. The agency should be able to resolve these issues by experimenting with the different choices for the window length and for linking the price level estimates for a new window to the estimates of the previous window.

Case 2: Price and Quantity Data are Available

Method 4: Apply the Section 4 Carry Forward Methodology

Little additional explanation is required here; just apply the methodology explained in section 3 to the elementary index context. Diewert, Fox and Schreyer (2018) have more details on how to apply the carry forward methodology for Paasche, Laspeyres, Fisher and Törnqvist indexes in the case of two observations.

Method 5: Apply the Weighted Time Product Dummy Multilateral Method

The basic economic model is still the price proportionality model defined by equations (19) above but now we assume that we have expenditure or quantity information on household purchases in addition to price information. With this extra information, it is preferable to take the economic importance of each commodity into account and replace the least squares minimization problem defined by (26) with the following weighted least squares minimization problem:[11]

$$(29) \min_{\rho,\beta} \{\Sigma_{t=1}^{T} \Sigma_{n\in S(t)} s_{tn}[\ln p_{tn} - \rho_t - \beta_n]^2\}$$

where the period t expenditure share on commodity n is $s_{tn} \equiv p_{tn}q_{tn}/p^t\cdot q^t$ for t = 1,...,T and $n\in S(t)$; see Diewert (2021c) for a discussion on the merits of different choices for the weights. Again, we need to make the normalization $\rho_1 = 0$ to obtain a unique solution ρ^* and β^* to (29). It turns out the solution will satisfy the following equations, which are the weighted counterparts to equations (27) and (28) (Diewert 2021c):

$$(30)\ \pi_t^* = \exp[\Sigma_{n\in S(t)} s_{tn}\ln(p_{tn}/\alpha_n^*)]\,; \qquad t = 1,\ldots,T;$$

$$(31)\ \alpha_n^* = \exp[\Sigma_{t\in S^*(n)} s_{tn}\ln(p_{tn}/\pi_t^*)/\Sigma_{t\in S^*(n)} s_{tn}]\,; \qquad n = 1,\ldots,N.$$

[11] Rao (1995) (2004) (2005; 574) was the first to consider this model using expenditure share weights. However, Balk (1980; 70) suggested this class of models much earlier using different weights. See also de Haan and Krsinich (2012) (2014) and Diewert and Fox (2018).

From (30) and (31), it can be seen that the period t estimated price level, π_t^*, is now a weighted geometric mean of the quality adjusted prices for products that are available in period t and the estimated quality adjustment factor for product n, α_n^*, is now a weighed geometric mean of all of the inflation adjusted prices for product n over all periods. Note that if the set of available products in periods r and t is the same, π_t^*/π_r^* will not collapse to a weighted Jevons index unless the expenditure shares in the two periods under consideration are equal.

Once the estimates for the π_t^* and α_n^* have been computed, we have two methods for constructing period by period aggregate price and quantity (or volume) levels, P^t and Q^t for t = 1,...,T. The way to see this is to consider the underlying equations (24) which were the equations $p_{tn} = \pi_t\alpha_n$ for t = 1,...,T and $n \in S(t)$. Take this equation for some n and t and multiply both sides of it by the observed quantity, q_{tn}, and sum the resulting equations. We obtain the following equations using the fact that $q_{tn} = p_{tn} \equiv 0$ for $n \notin S(t)$:

$$\begin{aligned}
(32)\ p^t \cdot q^t &= \Sigma_{n \in S(t)}\, p_{tn} q_{tn} && t = 1,\ldots,T \\
&= \pi_t \Sigma_{n \in S(t)}\, \alpha_n q_{tn} \\
&= \pi_t \Sigma_{n=1}^{N}\, \alpha_n q_{tn} && \text{since } q_{tn} = 0 \text{ if n does not belong to } S(t) \\
&= \pi_t \alpha \cdot q^t.
\end{aligned}$$

Because equations (24) will not hold exactly, with nonzero errors e_{tn}, equations (32) will only hold approximately. However, the approximate version of equations (32) allow us to form period t price and quantity aggregate levels, say P^t and Q^t, in two separate ways: the π_t^* estimates that are part of the solution to (29) can be used to form P^{t*} and Q^{t*} via equations (33) and the α_n^* estimates that are part of the solution to (29) can be used to form the aggregates P^{t**} and Q^{t**} via equations (34):

$$(33)\ P^{t*} \equiv \pi_t^*\ ; \quad Q^{t*} \equiv p^t \cdot q^t/\pi_t^*\ ; \qquad t = 1,\ldots,T;$$

$$(34)\ Q^{t**} \equiv \alpha^* \cdot q^t\ ;\ P^{t**} \equiv p^t \cdot q^t/\alpha^* \cdot q^t\ ; \qquad t = 1,\ldots,T.$$

Define the error terms $e_{tn} \equiv \ln p_{tn} - \ln\pi_t^* - \ln\alpha_n^*$ for t = 1,...,T and n = 1,...,N. If all $e_{tn} = 0$, then P^{t*} will equal P^{t**} and Q^{t*} will equal Q^{t**} for t = 1,...,T. However, if the error terms are not all equal to zero, then the statistical agency will have to decide on pragmatic grounds on which option to choose to form the aggregate price and quantity levels.[12]

It should be noted that $P^{t**} \equiv p^t\cdot q^t/\alpha^*\cdot q^t$ is a *quality adjusted unit value price level*.[13] There is also an inequality between P^{t*} and P^{t**} that is due to de Haan and Krsinich (2018; 763). From (30) and (33), we have $P^{t*} = \exp[\Sigma_{n\in S(t)}\, s_{tn}\ln(p_{tn}/\alpha_n^*)]$, which is a share weighted geometric mean of the period t quality adjusted prices, p_{tn}/α_n^*, for products that are actually present in period t. From (34), we have P^{t**} equal to the following expression:

$$
\begin{aligned}
(35)\; P^{t**} &\equiv p^t\cdot q^t/\alpha^*\cdot q^t \qquad\qquad t = 1,...,T \\
&= \Sigma_{n\in S(t)}\, p_{tn}q_{tn}/\Sigma_{n\in S(t)}\, \alpha_n^* q_{tn} \\
&= \Sigma_{n\in S(t)}\, p_{tn}q_{tn}/\Sigma_{n\in S(t)}\, \alpha_n^* (p_{tn})^{-1}p_{tn}q_{tn} \\
&= [\Sigma_{n\in S(t)}\, s_{tn}(p_{tn}/\alpha_n^*)^{-1}]^{-1} \\
&\le P^{t*}
\end{aligned}
$$

since a share weighted harmonic mean of the quality adjusted prices present in period t is always equal to or less than the corresponding share weighted geometric mean using Schlömilch's inequality (see Hardy, Littlewood and Polyá (1934; 26)). Note that $P^{t**} \le P^{t*}$ implies that $Q^{t**} \ge Q^{t*}$ for t = 1,...,T.

The axiomatic properties of the price levels π_t^* are studied in Diewert (2021c). They are reasonably good.

[12] De Haan and Krsinich (2018) were the first to realize that the results of a hedonic regression would lead to two separate ways to define the resulting aggregate price and quantity levels. See also Diewert (2020c) (2020d). If the accurate measurement of price levels is the target, then it is probably best to use P^{t*}; if the target is to measure aggregate quantity levels (and hence welfare), then it is probably best to use P^{t**}.

[13] The term "quality adjusted unit value price index" was introduced by Dalén (2001). Its properties were further studied by de Haan (2004) (2010), Silver (2010) (2011), de Haan and Krsinich (2018), Von Auer (2014) and Diewert (2020c) (2020d).

The issues of choosing a window length T for this multilateral method remain unresolved; statistical agencies can experiment with different choices for T. There is also the issue of linking the present window with the previous window.

From the viewpoint of the economic approach to index number theory, the use of this method should be confined to situations where the products in scope are close substitutes since the underlying economic assumption is that the products are perfect substitutes, except for random errors. Quality adjusted unit value price levels are appropriate in this situation but if the products are not close substitutes, it would be preferable to use the inflation adjusted carry forward prices methodology suggested by Diewert, Fox and Schreyer (2018) if the target index is a superlative index. Finally, Method 5 should not be used at higher levels of aggregation where substitution between elementary index categories may be low. At the second stage of aggregation it would be preferable to use Fisher, Walsh or Törnqvist indexes if actual price and quantity data are available or use pseudo Fisher indexes if the quantity data can only be approximated.

Method 6: The Use of Quality Adjusted Unit Value Price Levels

From the discussion of Method 5, it is clear that quality adjusted unit values can be used as price levels, provided that the commodities in scope for the elementary aggregate are close substitutes. However, it is not necessary to use the Weighted Time Product Dummy multilateral index number method in order to obtain estimates for the quality adjustment parameters, the components of the vector α. If the group of products under consideration consists of highly substitutable products and all of the products were purchased in the pre-lockdown period 0, then simply set α equal to p^0, the (unit value) price vector for the products in the pre-lockdown period. If all of the products were purchased for a number of pre-lockdown periods, say periods 0, −1, −2 and −3, and the price vectors for these periods were p^0, p^{-1}, p^{-2} and p^{-3}, then define α as follows:

$$(36)\ \alpha \equiv (1/4)[(p_{01})^{-1}p^0 + (p_{-1,1})^{-1}p^{-1} + (p_{-2,1})^{-1}p^{-2} + (p_{-3,1})^{-1}p^{-3}].$$

Thus α is set equal to the average of past pre-lockdown price vectors for the commodities in the group of commodities under consideration but these vectors of past prices are deflated by the price of the first commodity in order to eliminate the effects of general inflation between past periods for the group of commodities. The first commodity should be chosen to be the commodity with the largest average expenditure share in the group of commodities. If there are missing prices in the pre-lockdown periods, then instead of using the α defined by (36), the α defined by the Time Product Dummy multilateral method (Method 3 above) could be used to estimate the quality adjustment parameters.

From the viewpoint of the economic approach to index number theory, the use of quality adjusted unit values as estimates for price levels should only be applied if the commodities in the elementary group of commodities are close substitutes.[14]

Method 7: Linking Based on Relative Price and Quantity Similarity

A desirable property of the Fisher price index between two periods is the fact that the Fisher index will equal unity if prices in the two periods are equal even if the quantities demanded in the two periods are not equal. Most multilateral methods do not satisfy this strong identity test; they tend to satisfy a weaker identity test that says that the relative aggregate price levels between any two periods in the window of observations will equal unity provided that both prices *and* quantities are identical in the two periods being compared.

There is a recently developed multilateral method that satisfies the above strong identity test and can deal with missing observations. The method is based on building a set of Fisher index bilateral comparisons where each comparison is based on linking the periods

[14] It is possible to cluster N highly substitutable commodities in scope into quality groups based on their price per unit of a dominant characteristic. Group the N products into low quality, medium quality and high quality products based on their relative prices in the pre-lockdown period. Then aggregate price levels for each of the three groups of products could be constructed by simply taking unit values (without quality adjustment) for each group of products. We would end up with three elementary indexes in place of the single elementary index. Then these three separate indexes could be aggregated up into a single index using a superlative index number formula. This is feasible because we are assuming the availability of price and quantity data for Method 6. The advantage of this method is that it avoids the need for imputation.

that have the most *similar relative price structures*. Hill (2001) (2004) was an early pioneer in using this similarity of relative prices approach to multilateral index number theory in the time series context. The real time linking method described here is due to Diewert (2021c).

Initially, periods 1 and 2 are linked by the usual bilateral Fisher price index. When the data of period 3 become available, the price and quantity data of period 3 are linked to the corresponding data of either period 1 or 2, depending on which of these two periods has the most similar structure of relative prices. The bilateral Fisher index is used to link period 3 with period 1 if the measure of relative price similarity between periods 1 and 3 is higher than the measure of relative price similarity between periods 1 and 2. If the measure of relative price similarity between periods 2 and 3 is higher than the corresponding measure for comparing periods 1 and 3, then the bilateral Fisher index is used to link period 3 with period 2. When the data of period 4 become available, the data for period 4 are linked to the data of periods 1,2 or 3, depending on which of these 3 prior periods gives rise to the highest measure of price similarity. And so on. In practice, measures of *relative price dissimilarity* are used to link the data of two periods, using the lowest measure of dissimilarity to do the linking. At the first stage of the network of comparisons, the two periods that have the most similar structure of relative prices is chosen. At the next stage of the comparison, look for a third period that had the most similar relative price structure to the first two periods and link in this third country to the comparisons of volume between the first two countries and so on.

A key aspect of this linking methodology is the choice of the measure of similarity (or dissimilarity) of the relative price structures of two countries. Various measures of the similarity or dissimilarity of relative price structures have been proposed by Allen and Diewert (1981), Kravis, Heston and Summers (1982; 104-106), Hill (1997) (2001) (2004) (2009), Aten and Heston (2009) and Diewert (2009). The predicted share dissimilarity measure recently proposed by Diewert (2021c) seems to be the most promising but the method needs to be more thoroughly tested before it can be suggested to statistical agencies for general use. A major advantage of this new method of linking periods is that the strong identity test will always be satisfied; i.e., if prices in the current period are the same as the prices in a past period, the estimated price levels pertaining to these two periods will always be identical even if quantities or expenditures are not identical. If the prices in the current period are proportional to the prices in a prior period, then the ratio of the current period price level to the prior period price level will be equal to the factor of proportionality. Another advantage of Diewert's method is that it is not necessary to choose a window length. There can never be a chain drift problem using this new multilateral method.

6. Other Measurement Problems

6.1. No Agency Employee Price Collection

Most statistical agencies stopped sending employees to retail outlets to collect prices during pandemic periods. Some agencies have switched to web scraping; i.e., they collect online prices over the internet. The collected prices will not be perfectly comparable with the previously collected in store prices. Cavallo (2017) did a large scale comparison of in store prices versus online prices (excluding transport costs) across 10 countries and found little difference between in store and online prices; online prices over the comparable in-store prices were on average 4% lower. The average markup ranged from −13% for Japan to +5% for Australia. See also Cavallo (2013) and Cavallo and Rigobon (2016). These results provide some justification for comparing a web scraped price for a specific product with a collected price for the same product in a prior period. Under lockdown conditions, home delivery of products purchased online increased dramatically. On the other hand, household travel expenses decreased due to fewer in store shopping trips. As these travel expenses are in scope for household expenditures, it may make sense to collect online prices that include delivery since the delivered price is the price that the consumer actually faces for the product. The higher price for the delivered product will be offset by lower household transportation costs. In general, we endorse the collection of web scraped data to replace previous data that were collected by agency employees. However, some care should be taken to not collect online prices for goods or services which were never actually consumed by any household. Examples of such services are be listed airline fares or listed prepaid holiday packages that are eventually cancelled. (How exactly should cancellation fees be treated in a CPI?)

6.2 Lack of Information on Current Household Expenditure Weights

It will be very difficult for statistical agencies to find current period expenditure share or quantity weights for their elementary index categories. The problem is that the "representative" basket for each month is changing rapidly as the virus spreads and lockdown rules change to react to current conditions. Here are some possible ways for NSOs to obtain current information on household expenditures:

- Some countries (such as the US and the UK) have continuous household expenditure surveys. Usually, the sample size for such surveys is small so, for example, the US Bureau of Labor Statistics Consumer Expenditure Survey does not have a big enough sample size to allow monthly publication of the implied monthly weights. It publishes semi-annual estimates. The way forward here is to increase the sample size. For countries that currently do not have a continuous consumer expenditure survey, it is recommended that they start one. National governments will have to allocate extra resources to fund a continuous survey.
- Some private companies collect consumer expenditure data (along with prices and quantities) on a continuous basis for a sample of households using scanner data. NSOs can purchase these data (at a fair price) or set up their own competing company if they are unable to establish a satisfactory consumer expenditure survey.
- National governments can appeal to their business communities to persuade large firms producing consumer products to donate their electronic data to the NSO. Many countries, including Canada, have a Statistics Act which can be used to compel firms and households to provide information to NSOs. However, in general, NSOs are reluctant to use compulsion in order to obtain data. Many large retailers around the world are already donating their data and it should be possible

for more firms to be persuaded to do this. This information will help to produce a better CPI and it will also allow much better production accounts to be produced.

- Credit card companies collect information on household purchases of consumer goods and services. If the expenditure information could also be combined with product codes, this information would enable the construction of consumer price indexes by location and demographic group. For some countries, it may be possible to access this information source. For other countries, it may not be possible for the statistical agency to access this information due to privacy concerns. See Carvalho et al. (2020) and Dunn, Hood, and Driessen (2020) for examples of how such information can be used to analyze changes in expenditure patterns.

6.3 Should the CPI be Revised?

From section 6.2, it can be seen that NSOs will not be able to produce very accurate period t basket updates q^{bt} that approximate actual period t consumption q^t in a timely fashion (if they can produce them at all). However, in time, better estimates for actual consumption in past periods may become available. Smoothing a sample of collected monthly household expenditures (by taking a moving average for example) will probably lead to more accurate trend estimates for monthly household expenditures, but the trend can only be calculated after some months have passed. The question then arises: should the CPI be revised in the light of improved information that becomes available after the release date? From a statistical point of view, the answer to this question is yes. However, for many countries, a monthly CPI must be provided to the public and no revisions are allowed.

Scanner data along with the usual information on retail sales can be massaged to produce some rough and ready weights in real time.[15] NSOs will simply have to announce that their new estimates for inflation and economic growth are only very approximate

[15] Scanner data from retail outlets is not perfectly well suited for a CPI: retail outlets sell to tourists, foreign firms, governments and to domestic firms as well as to households. Thus scanner data collected directly from households is preferable.

estimates. A country's national accounts are allowed to be revised and this revision process is generally accepted by the public, hence estimates of economic growth can be revised. This is not the case for the CPI. A country could produce at least two CPIs: one that is not revised and is based on available information at the month of production of the index and another that is allowed to be revised in the light of information that becomes available at a later date.

This strategy has been successfully used by the Bureau of Labor Statistics in the US where two indexes are released at the same time; the first one ("CPI-U") is not revisable and the second one ("C-CPI-U") is allowed to be revised (and approximates a superlative index after the last revision). The second CPI can be labeled as an analytic CPI and can be used by economic analysts who require more accurate historical information on inflation. The first type of traditional CPI produced under lockdown conditions will necessarily be much more inaccurate; it will be very difficult to obtain adequate approximations to actual consumption during the start of the lockdown period due to the absence of accurate survey information on consumer expenditures. Users need to be alerted to this problem.

In section 6.2 we attempted to anticipate the problems that many statistical agencies will face in trying to update their baskets to reflect the lockdown realities. We realize that new lockdown baskets will not be available to many, if not most, NSOs. Our conclusion boils down to this: if later information shows that the early lockdown indexes are very inaccurate, then set the current CPI price level to the best estimate possible even if it is necessary to use a different methodology than was used in the pre-lockdown periods. For the revisable CPI, new information should be used to revise previous indexes.

6.4 The Stockpiling Problem

Lockdowns have led governments to limit trips to retail outlets for purchases of food and other essential goods such as pharmaceutical products. These regulations plus the reactions of households to cut down on their shopping trips to limit the risk of infection have led households to accumulate large *stockpiles* of essential storable goods. Thus at the initial stages of a lockdown, there were large increase in purchases of storable goods but actual consumption of these goods was less. In other words, it becomes necessary to distinguish actual household *consumption* of storable goods from the *acquisition* of the goods. In principle, the national statistical agency will have to decide between these two approaches to the production of a CPI. The acquisitions approach is of course much more practical. In order to implement an actual consumption approach, the NSO would require a household inventory survey which would be costly.

From a welfare point of view, it is monthly consumption of goods and services which is most relevant but it will generally be more convenient to stick to an acquisitions approach to the measurement of consumption. If the actual consumption approach to the scope of the CPI is chosen, then in principle, the stocks of storable items need to be measured at the beginning and end of each period.[16] If the acquisitions approach to storable goods is taken, then household purchases of essential storable goods at the beginning of the lockdown period will be very much larger than pre-lockdown purchases of the same goods. Once the lockdown has been in place for a month or two, then purchases of storables should fall back to pre-lockdown levels. But the problem here is that the assumption of a constant basket equal to a pre-lockdown basket for all post lockdown periods may be a rather poor assumption.

6.5 How Should Scanner Data be Combined with Web Scraped Data?

Many statistical agencies now have access to scanner data from some retailers. How exactly should the indexes which are generated by the use of these data be combined with

[16] Real actual consumption of a storable good is equal to beginning of the period inventory stock plus new purchases of the good less end of period stock of the good. In principle, if actual consumption is the target concept, then household stocks of storables should be capitalized and added to household wealth. In normal times, the services provided by these storable stocks probably should not be added to the current flow of consumption unless one argues that these stocks are desirable in their own right as a form of insurance against future supply shocks. In times of a pandemic, such an argument seems reasonable. Note that not recognizing a flow of services from the storables stock is a different treatment from the treatment of the services that consumer durables provide over their useful lifetime. Stocks of consumer durables should also be capitalized and added to household wealth but the services that durables render during a month need to be recognized as part of actual consumption.

traditional price data collected by statistical agency employees or by the use of web scraped data?

In general, it is preferable if the contribution of these two sources of price data be combined in an index which weights the prices according to their economic importance; i.e., to their shares of expenditure in the elementary category under consideration. It is not a problem to calculate expenditure shares for the scanner data but the web scraped data will not come with the associated expenditure data and so weighting the two sources of data by their relative quantities or expenditure shares will not be possible. In the end, some rough explicit or implicit estimate of the relative economic importance of the two sources of data will have to be made. Area specialists in NSOs will have to provide approximate weights for each elementary category that uses the two sources of information.

7. Conclusion

We suggest that statistical offices concentrate on getting more up to date expenditure weights for the post-lockdown period so that inflation during the lockdown period can be more accurately measured. When the lockdown ends, we suggest the use of a Fisher index, linking the first post-lockdown period to the last pre-lockdown period.

We have shown how the use of inflation adjusted carry forward prices for missing products, as recommended by international agencies, will typically lead to an understatement of inflation using our concept of market clearing imputed prices for missing products. However, these imputed prices require econometric estimation and so implementing this approach can only be done on a retrospective basis using econometric techniques.

Three steps that NSOs can take to provide as much information as possible on price indexes during a time of lockdown are:

1. Collect whatever prices are available, including from non-traditional sources. For missing prices, use inflation adjusted carry forward prices. While we favour using reservation prices, we acknowledge that currently it is unlikely that NSOs will be able to estimate these in a timely fashion.
2. Start a program to obtain current expenditure weights for the consumption basket.
3. Produce a revisable CPI as an analytical series that can be updated as new methodology is developed and new data sources are exploited. Statistics Canada has followed this advice.

The U.S. Bureau of Labor Statistics' (2020) approach[17] to dealing with the pandemic is very much in line with the approach advocated in this paper; i.e., the BLS produces a headline non-revisable standard fixed basket Lowe index while at the same time, it produces an approximation to a Törnqvist index which is improved over a two year revision period. Thus this supplementary index eventually measures inflation using weights that reflect current consumer expenditure patterns. Given that lockdown conditions have applied in varying degrees in many countries for many months, it is important to have information on current period household expenditure patterns so that meaningful estimates of consumer price inflation can be produced during the lockdown periods.

Finally, it is unlikely that expenditure patterns will revert to the pattern that prevailed in periods just before the first lockdown period. This reinforces the case for obtaining more current estimates for household expenditures by category.

References

Aizcorbe, A., C. Corrado and M. Doms (2000), "Constructing Price and Quantity Indexes for High Technology Goods", Industrial Output Section, Division of Research and Statistics, Board of Governors of the Federal Reserve System, Washington DC.

Allen, R.C. and W.E. Diewert (1981), "Direct versus Implicit Superlative Index Number Formulae", *Review of Economics and Statistics* 63, 430-435

[17] See Appendix A in Diewert and Fox (2020) and the references there for more materials on the BLS approach.

Aten, B. and A. Heston (2009), “Chaining Methods for International Real Product and Purchasing Power Comparisons: Issues and Alternatives”, pp. 245-273 in *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, D.S. Prasada Rao (ed.), Cheltenham UK: Edward Elgar.

Australian Bureau of Statistics (2017), “An Implementation Plan to Maximise the Use of Transactions Data in the CPI”, Information Paper 6401.0.60.004, 16 June, Canberra.

Balk, B.M. (1980), “A Method for Constructing Price Indices for Seasonal Commodities”, *Journal of the Royal Statistical Society, Series A* 143, 68-75.

Bureau of Labor Statistics (2020), “Effects of COVID-19 Pandemic on BLS Price Indexes”, April 7, 2020. Web address: https://www.bls.gov/bls/effects-of-covid-19-pandemic-on-bls-price-indexes.htm

Carvalho, V.M., J.R. Garcia, S. Hansen, Á. Ortiz, T. Rodrigo, J.V. Rodríguez Mora and J. Ruiz (2020), “Tracking the COVID-19 Crisis with High-Resolution Transaction Data”, Cambridge-INET Working Paper Series No: 2020/16, University of Cambridge.

Cavallo, A. (2013), “Online and Official Price Indexes: Measuring Argentina's Inflation”, *Journal of Monetary Economics* 60:2, 152–165.

Cavallo, A. (2017), “Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers”, *American Economic Review* 107:1, 283–303.

Cavallo, A. (2020), “Inflation with Covid Consumption Baskets,” NBER Working Paper 27352, Cambridge, MA.

Cavallo, A. and R. Rigobon (2016), “The Billion Prices Project: Using Online Prices for Inflation Measurement and Research”, *Journal of Economic Perspectives* 30:2, 151–178.

Court, A.T. (1939), “Hedonic Price Indexes with Automotive Examples”, pp. 99-117 in The *Dynamics of Automobile Demand*, New York: General Motors Corporation.

Dalén, J. (2001), “Statistical Targets for Price Indexes in Dynamic Universes,” Paper presented at the Sixth meeting of the Ottawa Group, April 2-6, Canberra, 2001.

Deaton, A. and W.E. Diewert (2002), “Conceptual Foundations for Price and Cost of Living Indexes”, Chapter 2 (pp. 38-93) in *At What Price? Conceptualizing and Measuring Cost of Living and Price Indexes*, C.L. Schultze and C. Mackie (eds.), Washington, DC: National Academy Press.

de Haan, J. (2004), “Estimating Quality-Adjusted Unit Value Indices: Evidence from Scanner Data,” Paper presented at the Seventh EMG Workshop, Sydney, Australia, December 12–14.

de Haan, J. (2010), “Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and Re-pricing Methods”, *Jahrbücher für Nationökonomie und Statistik* 230, 772-791.

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 Workshop 2012, University of New South CourtWales, 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.

de Haan, J. and F. Krsinich (2018), "Time Dummy Hedonic and Quality Adjusted Unit Value Indexes: Do They Really Differ?", *Review of Income and Wealth* 64:4, 757-776.

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.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 114-145.

Diewert, W.E. (1978), "Superlative Index Numbers and Consistency in Aggregation", *Econometrica* 46, 883-900.

Diewert, W.E. (1992), "Fisher Ideal Output, Input and Productivity Indexes Revisited", *Journal of Productivity Analysis* 3, 211-248.

Diewert, W.E. (2003), "Hedonic Regressions: A Review of Some Unresolved Issues", Paper presented at the Seventh Meeting of the Ottawa Group, Paris, 27–29 May.

Diewert, W.E. (2009), "Similarity Indexes and Criteria for Spatial Linking", pp. 183-216 in *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, D.S. Prasada Rao (ed.), Cheltenham, UK: Edward Elgar.

Diewert, W.E. (2021a), "The Economic Approach to Index Number Theory", Draft Chapter 5 in Consumer Price Index Theory, Washington D.C.: International Monetary Fund. https://www.imf.org/en/Data/Statistics/cpi-manual.

Diewert, W.E. (2021b), "Elementary Indexes", Draft Chapter 6 in Consumer Price Index Theory, Washington D.C.: International Monetary Fund. https://www.imf.org/en/Data/Statistics/cpi-manual.

Diewert, W.E. (2021c), "The Chain Drift Problem and Multilateral Indexes", Draft Chapter 7 in Consumer Price Index Theory, Washington D.C.: International Monetary Fund. https://www.imf.org/en/Data/Statistics/cpi-manual.

Diewert, W.E. (2021d), "Quality Adjustment Methods", Draft Chapter 8 in Consumer Price Index Theory, Washington D.C.: International Monetary Fund. https://www.imf.org/en/Data/Statistics/cpi-manual.

Diewert, W.E., J. de Haan and R. Hendriks (2015), "Hedonic Regressions and the Decomposition of a House Price index into Land and Structure Components", *Econometric Reviews*, 34:1-2, 106-126.

Diewert, W.E. and R. Feenstra (2019), "Estimating the Benefits of New Products: Some Approximations", Discussion Paper 19-02, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E. and K.J. Fox (2018), "Substitution Bias in Multilateral Methods for CPI Construction Using Scanner Data," UNSW Business School Research Paper No. 2018-13. Available at SSRN: https://ssrn.com/abstract=3276457

Diewert, W.E. and K.J. Fox (2020), "Measuring Real Consumption and CPI Bias under Lockdown Conditions", NBER Working Paper 27144, Cambridge, MA. https://www.nber.org/papers/w27144

Diewert, W.E., K.J. Fox and P. Schreyer (2018), "The Digital Economy, New Products and Consumer Welfare", Economic Statistics Centre of Excellence (ESCoE) Discussion Paper 2018-16, London, UK.

Dunn, A., K. Hood and A. Driessen (2020), "Measuring the Effects of the COVID-19 Pandemic on Consumer Spending using Card Transaction Data," U.S. Bureau of Economic Analysis Working Paper WP2020-5.

Dutot, Charles, (1738), *Réflexions politiques sur les finances et le commerce*, Volume 1, La Haye: Les frères Vaillant et N. Prevost.

Eurostat (2020), "Guidance on the Compilation of the HICP in the Context of the COVID-19 Crisis", Methodological Note, Directorate C, Unit C4, April 3. Web Address: https://ec.europa.eu/eurostat/documents/10186/10693286/HICP_guidance.pdf

Feenstra, R.C. (1994), "New Product Varieties and the Measurement of International Prices", *American Economic Review* 84:1, 157-177.

Fisher, I. (1922), *The Making of Index Numbers*, Boston: Houghton-Mifflin.

Hardy, G.H., J.E. Littlewood and G. Pólya (1934), *Inequalities*, Cambridge: Cambridge University Press.

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.

Hicks, J.R. (1940), "The Valuation of the Social Income", *Economica* 7, 105–124.

Hill, R.J. (1997), "A Taxonomy of Multilateral Methods for Making International Comparisons of Prices and Quantities", *Review of Income and Wealth* 43(1), 49-69.

Hill, R.J. (2001), "Measuring Inflation and Growth Using Spanning Trees", *International Economic Review* 42, 167-185.

Hill, R.J. (2004), "Constructing Price Indexes Across Space and Time: The Case of the European Union", *American Economic Review* 94, 1379-1410.

Hill, R.J. (2009), "Comparing Per Capita Income Levels Across Countries Using Spanning Trees: Robustness, Prior Restrictions, Hybrids and Hierarchies", pp. 217-244 in *Purchasing Power Parities of Currencies: Recent Advances in Methods and Applications*, D.S. Prasada Rao (ed.), Cheltenham UK: Edward Elgar.

Hofsten, E. von (1952), *Price Indexes and Quality Change*, London: George Allen and Unwin.

ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2004), *Consumer Price Index Manual: Theory and Practice*, Peter Hill (ed.), Geneva: International Labour Office.

IMF (2020), "Consumer Price Index Continuity", Unit C-4, Price Statistics, April. Web Address: https://statswiki.unece.org/display/CCD2/Compilation+of+CPI+in+times+of+COVID-19?preview=/278037166/279773977/STA%20CPI%20Business%20Continuity.pdf

Intersecretariat Working Group on Price Statistics (2020), "Consumer Price Index: Business Continuity Guidance", online at: https://statswiki.unece.org/display/CCD2/Compilation+of+CPI+in+times+of+COVID-19?preview=/278037166/279776928/IWGPS%20CPI%20Continuity%20Note.pdf

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.

Jevons, W.S., (1865), "The Variation of Prices and the Value of the Currency since 1782", *Journal of the Statistical Society of London* 28, 294-320; reprinted in *Investigations in Currency and Finance* (1884), London: Macmillan and Co., 119-150.

Konüs, A.A. (1924), "The Problem of the True Index of the Cost of Living", translated in *Econometrica* 7, (1939), 10-29.

Kravis, I.B., A. Heston and R. Summers (1982), *World Product and Income: International Comparisons of Real Gross Product*, Statistical Office of the United Nations and the World Bank, Baltimore: The Johns Hopkins University Press.

Pollak, R.A. (1983), "The Theory of the Cost-of-Living Index", pp. 87-161 in *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada

Rao, D.S. Prasada (1995), "On the Equivalence of the Generalized Country-Product-Dummy (CPD) Method and the Rao-System for Multilateral Comparisons", Working Paper No. 5, Centre for International Comparisons, University of Pennsylvania, Philadelphia.

Rao, D.S. Prasada (2004), "The Country-Product-Dummy Method: A Stochastic Approach to the Computation of Purchasing Power parities in the ICP", paper presented at the SSHRC

Conference on Index Numbers and Productivity Measurement, June 30-July 3, 2004, Vancouver, Canada.

Rao, D.S. Prasada (2005), "On the Equivalence of the Weighted Country Product Dummy (CPD) Method and the Rao System for Multilateral Price Comparisons", *Review of Income and Wealth* 51:4, 571-580.

Silver, M. (2010), "The Wrong and Rights of Unit Value Indices", *Review of Income and Wealth* 56:S1, 206-206.

Silver, M. (2011), "An Index Number Formula Problem: the Aggregation of Broadly Comparable Items, *Journal of Official Statistics* 27:4, 1–17.

Summers, R. (1973), "International Comparisons with Incomplete Data", *Review of Income and Wealth* 29:1, 1-16.

Theil, H. (1967), *Economics and Information Theory*, Amsterdam: North-Holland.

UNECE (2020) (A. Bredt and C. Boldsen), "Compilation of CPI in Times of COVID-19", web address: https://statswiki.unece.org/display/CCD2/Compilation+of+CPI+in+times+of+COVID-19

von Auer, L. (2014), "The Generalized Unit Value Index Family", *Review of Income and Wealth* 60, 843-861.

Walsh, C.M. (1901), *The Measurement of General Exchange Value*, New York: Macmillan and Co.