Scanner Data, Chain Drift, Superlative Price Indices and the Redding-Weinstein CES Common Varieties Price Index

Naohito Abe
Institute of Economic Research, Hitotsubashi University
Naka, Kunitachi, Tokyo Japan 186-8603,
nabe@ier.hit-u.ac.jp
Phone: 86-42-580-8347, Fax: 86-42-580-8333

D.S. Prasada Rao
The University of Queensland and the Hitotsubashi Institute for Advanced Study (HIAS)

Abstract

Scanner data are being increasingly integrated by national statistical offices into the compilation of the consumer price index. An appealing feature of scanner data is the availability of quantity data along with price quotations recorded at the point of sale which makes it possible to consider the full range of index number methods at the elementary level. Chain drift is a serious problem encountered in the application of superlative indexes like the Fisher and Tornqvist which is often resolved through the use of the Gini-Elteto-Koves-Szulc and the Geary-Khamis methods. These somewhat ad hoc solutions lack theoretical foundations and interpretation as cost of living indexes. In this paper we establish credentials of the exact *CES common varieties (CCV)* price index proposed in Redding and Weinstein (2020) by proving that the index is transitive and therefore eliminates chain drift when using high frequency data. We demonstrate the effectiveness of the index by applying the CCV index along with a raft of other indexes to Japanese scanner data. Empirical results suggest a bias associated with the use of Chained Sato-Vartia index is 5.89 percent at annual rate. The paper offers additional insights. First, we show that, unlike the Sato-Vartia index, the CCV index is monotonic with respect to current prices. Second, the implicit quantity index based on the CCV price index is a measure of welfare change. Finally, the paper resolves the problem of specification of normalization condition in Redding and Weinstein (2020) by providing a necessary and sufficient condition for the index to satisfy the commensurability property which ensures that the index is independent of units of measurement. The CCV index with these demonstrated properties may be considered superior to the Fisher ideal index. The paper also discusses some of the practical issues surrounding the presence of stocking behaviour, similar to that observed during the COVID-19 pandemic, and the application of the CCV price indexes.

Key words: Price comparisons; Preference Heterogeneity; Logarithmic Indices; Transitivity; Scanner Data; Chain Drift
JEL Codes: E31, C43

23 April, 2022

The authors gratefully acknowledge helpful suggestions and comments from Robert Feenstra, Stephen Redding and David Weinstein on an earlier incarnation of this paper. Thanks are also due to Ryo Jinnai, Chiaki Moriguchi and Yuta Takahashi for their comments on this version. Abe's work was supported by JSPS KAKENHI, 19H01467, 16H06322, and 18H00864. Rao acknowledges support from Hitotsubashi Institute for Advanced Study (HIAS) which facilitated visits to HIAS and funding support through the Australian Research Council Discovery Project DP170103559.

1. Introduction

Recently, more and more researchers and statistical offices in various countries are incorporating high frequency point of sale scanner data into compilation of aggregate measures of price change. Often they encounter the problem of chain drift when chained price index numbers are used. Diewert (2020, p.3) describes: "…*Chain drift* occurs when an index does not return to unity when prices in the current period return to their levels in the base period". The seriousness of chain drift problem is well documented in Diewert (2018, 2020)[1]. The spike in purchases and then return to the original levels is a major driver of the chain drift problem. Feenstra and Shapiro (2003) considered the chain drift problem caused by sales and stocking behaviour and suggested the use of fixed base indices as a solution. However fixed base index numbers have problem including sensitivity of price change measures to the choice of the base period.

The cost-of-living indexes (COLIs) like the Fisher, Tornqvist and S-V indices exhibit significant chain drift when using scanner data (de Haan and van der Grient, 2011). To date, solutions to this problem rely heavily on multilateral index number formulae which satisfy transitivity. The use of Gini-Elteto-Koves-Szulc (GEKS) (Gini, 1931; Elteto and Koves, 1964; Szulc, 1964) index and the rolling-window GEKS methods (Ivancic, Diewert and Fox, 2009 and 2011; de Haan, 2008; de Haan and Krsinich, 2014; van Auer, 2019) and variants of the Geary (1958) and Khamis (1972) method have been proposed and being explored (Chessa, 2016; Lamboray, 2017). For example, the GEKS method builds on the Fisher binary index (Diewert, 1976) but has limited economic theoretic interpretation (Neary, 2004). The GEKS method at best can be considered as a technique that generates transitive comparisons which deviate the least from a set of non-transitive bilateral comparisons. These methods offer viable solutions to the problem lack but solid economic theoretic foundations and remain essentially heuristic.

One main reason for the observed chain drift is that while the Fisher, Tornqvist and SV indices are known to be exact (Diewert, 1976, Sato, 1976 and Feenstra, 1994) for Konus (1924) based COLI measure of price change from the base to a comparison period is defined as the ratio of minimum expenditures required to attain a given utility level at prices prevailing in respective periods[2], these indices do not satisfy transitivity.

Our research reported here began when we observed the following puzzle with logarithmic price index numbers currently in use as deflators. By definition, Konus-based COLI must be monotonically increasing in prices in comparison period[3] and transitive, i.e., price change from period *s* to *t* multiplied by price change from period *t* to *u* must equal price change from *s* to *u*.[4] But two of the well-known logarithmic indices, the Tornqvist (1931) and the Sato (1976)-Vartia (1976) (S-V from here on) indices, which are the COLI, respectively for homothetic translog and constant elasticity of substitution (CES) utility functions (Diewert, 1976; Sato, 1976; and Feenstra, 1994), fail to satisfy monotonicity (Reinsdorf and Dorfman ;1999) and transitivity (can be seen from the numerical example presented in Section 3.2).[5] The Fisher index, while it satisfies montonicity fails to satisfy transitivity.

---

[1] Diewert (2020) will be included as Chapter 7 in the latest version of the *CPI Manual* (ILO/IMF/OECD/UNECE/Eurostat/The World Bank, 2020).

[2] If the minimum expenditure necessary to attain utility level U at prices, p, is denoted by $E(p,U)$, then the Konus COLI for measuring price changes is given by $E(p_t,U)/E(p_s,U)$. See section 2 for more details.

[3] This follows from properties of expenditure function from duality theory of consumer behavior.

[4] Since COLI is the ratio $E(p_t,U)/E(p_s,U)$, it should be automatically transitive.

[5] Non-transitivity of S-V index and the process of generating transitive indices from S-V framework are reported in Abe and Rao (2019).

Before we get further on the puzzle, we note that monotonicity of an aggregate price change measure is intuitive whereas importance of transitivity is less obvious. We refer to the classic Samuelson and Swamy (1974) (S-S) work where role of transitivity is discussed in several parts of the paper. S-S conclude that if the price index for deflating nominal expenditures does not possess transitivity, the implied real consumption violates transitivity in consumer preferences, which is reinforced in the quote:

"*Conclusion: So long as we stick to the economic theory of index numbers, the circular test is as required as is the property of transitivity itself. And this regardless of homotheticity or nonhomotheticity."* (Samuelson and Swamy, 1974, page, 576).

We note in passing that commonly used formulae[6] such as the Laspeyres, Paasche, Fisher, Tornqvist, S-V indices, and their chained counterparts fail transitivity. The observation by Samuelson-Swamy is further reinforced by the fact that transitivity helps mitigate the chain-drift effect often observed in applications involving scanner data.

In this paper, we offer an explanation of the puzzle by drawing on the path breaking and influential research[7] reported in Redding and Weinstein (2020) where R-W propose a class of aggregate price indices with taste shocks. We focus in particular on the insightful observation of R-W that the observed price and expenditure data do not usually match the implied levels from an assumed preference structure resulting in demand residuals. One of the main contributions of R-W is that the S-V index, which relies on observed prices and expenditure shares, fails to properly deal with demand residuals and therefore is prone to bias.

We find demand residuals discussed in R-W to be the key that resolves our puzzle as to why COLI such as Tornqvist and S-V indices fail monotonicity and transitivity. Our contribution here is that we are able to prove that the introduction of taste shocks (in R-W, 2020) into the utility function leads to a COLI that not only eliminates bias in aggregate price change measure (which was the main concern of R-W) but also results in an index that satisfies both monotonicity and transitivity. We employ the same R-W normalization that the geometric mean of taste shocks is a constant. Apart from the intuitive justification that this normalization ensures that logarithm of demand shocks has a mean equal to zero, we establish, in this paper, that this normalization belongs to the class of normalizations that ensure that the resulting measure of price change is independent of units of measurement. In this process we characterize the class of normalizations that ensure commensurability.

In this paper we take a broader view of the "taste" shocks in R-W (2020). There can be several sources of shocks that may ultimately result in demand function residuals, such as taste shocks, quality changes (R-W, 2020)[8]; stocking behavior (Hendel and Nev, 2006); habit formation (Dynan, 2000) seasonal effects (Osbone;1988); and also pure random shocks. In this paper we consider the all-inclusive label, *demand shocks*.

An equally important contribution of this paper is to demonstrate that the general class of logarithmic price aggregators proposed by R-W (2020) has the potential to bridge the existing gap between consumer theory and practical measurement of price changes using index number formulae. We primarily focus on the common goods version of the R-W CES Unified Price Index (CUPI) ignoring the variety effects though the

---

[6] See Diewert (1983) for an account of history of index numbers and for algebraic expressions of these indices.
[7] See Crawford and Neary (2019); Ehrilch et al., (2019); Demuynck and Parenti (2018); Feenstra, Xu and Antonaides (2018); and Martin (2019a and 2019b). These studies refer to the NBER Working Paper Series, Redding and Weinstein (2018).
[8] R-W (2020) at various places acknowledge that taste shocks may represent quality changes and also referred to these shocks as demand shocks.

results we establish here hold for the case with variety effects[9]. Like R-W, given its tractability and importance of CES in consumer and producer behavior theory; macroeconomics; industrial organization; and international trade, we focus on CES preferences. Consequently, we consider CUPI as a prime example of generalized logarithmic index with demand shocks.

The paper makes several important contributions. First, we establish two important theoretical properties, monotonicity and transitivity, of CUPI thus satisfying an important requirement of Samuelson and Swamy (1974) for a price index number to result in welfare comparisons. Second, we prove that the indirect quantity index derived using CUPI gives a measure of welfare change. Third, on the practical side, we argue that CUPI, due to its transitivity property, is well suited for measuring aggregate price changes with high frequency scanner data as it eliminates chain drift[10], a problem frequently encountered with other price indices (see Diewert, 2018, 2020). Fourth, a distinct practical advantage of CUPI over other transitive methods used to reduce chain drift (Diewert, 2020) is that there is no need to revise estimates of price changes when data for a new period is introduced. Finally, in a significant finding we offer a closure to an open question in R-W (2020, Section II), choice of normalization condition on taste parameters. Invoking a fundamental property expected of any measure of aggregate price change, independence of units of measurement (commensurability), we are able to prove that the geometric mean normalization of taste parameters that underpins CUPI, belong to the class of normalizations that ensure y commensurability.

The paper is organized as follows. Section 2 presents notation and basic demand theoretic results for CES with time varying demand shocks. The CUPI is also described. Section 3 forms the core of the paper with results on transitivity and monotonicity, Section 4 focuses on normalization of demand shock parameters. Section 5 reports empirical results from the analysis of Japanese scanner data. Section 6 concludes demonstrating CUPI provides a bridge between theory and practice of measuring aggregate price changes over time.

2. Notation, the Sato-Vartia Index and the CES Unified Price Index (CUPI)

2.1 Notation

Let prices, quantities, and time-varying demand shock parameters for N (> 1) commodities at time t be, respectively, denoted by the vectors

$$p_t = \left(p_{1t}, p_{2t}, ..., p_{Nt}\right),\ \ q_t = \left(q_{1t}, q_{2t}, ..., q_{Nt}\right), \text{ and } \varphi_t = \left(\varphi_{1t}, \varphi_{2t}, ..., \varphi_{Nt}\right).$$

While most resutls discussed in this paper can be obtained with the case with variable variety[11], to keep the discussion simple, we assume that the number of commodities, N, is the same for all time periods.

The demand shock parameters, φ_{it}, are essentially the taste shock parameters discussed in R-W (2020). Demand shock parameters here can represent taste shocks, quality change, stocking behaviour in purchases as well as random shocks. We make use of R-W (2020) specification of the following homothetic CES with time varying shock parameters.

[9] R-W refer to the index with common varieties as the CES Common Varieties index. We prefer to use the term CES Unified Price Index as it captures the essence of the R-W contribution.
[10] We elaborate further on chain drift in Section 4.
[11] The R-W framework includes variety effects with different number of commodities in each period. Our results can be extended to include variety effects easily.

$$U_t\left(q_t;\varphi_t,\sigma\right)=\left(\sum_{i=1}^{N}\left(\varphi_{it}q_{it}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}},\sigma>1 \tag{1}$$

where σ *is* the elasticity of substitution.

2.2 CES Utility function and key results

We consider a representative consumer who solves the following utilty maximization problem:

$$\text{Max } U_t\left(q_t;\varphi_t,\sigma\right) \quad \textit{such that } \sum_{i=1}^{N}p_{it}q_{it}=Y_t. \tag{2}$$

where Y_t exogeneous income at time t.

Unit expenditure function: The expenditure function, $E(p_t,U_t)$, minimum expenditure necessary to attain utility level U_t at prices p_t, is the product of the unit expenditure function, $C\left(p_t;\varphi_t,\sigma\right)$, and the utility as follows,

$$E\left(p_t,U_t\right)=C\left(p_t;\varphi_t,\sigma\right)\times U_t=\left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}\times U_t. \tag{3}$$

Normalization of taste parameters: The CES utility function in (1) represents the same preference relation and relative utilities when all the demand shock parameters are multiplied by a positive constant. In order to determine the demand shocks uniquely, R-W impose the following normalization on φ_i's:

$$\prod_{i=1}^{N}\left(\varphi_{it}\right)=\varphi. \tag{4}$$

Although the above condition looks natural and determines these parameters uniquely, other normalizations based on arithmetic means, or any type of conditions that help uniquely determine the demand shock parameters are possible. There arises a problem of choice between normalizations. R-W prefer normalization in (4) but discuss robustness of their findings when alternative specifications (R-W, 2020, p. 518) are used. We return to this problem in Section 4 where we prove that the geometric normalization in (4) is the only meaningful normalization if aggregate measures of price change are to be independent of units of measurement.

Marshallian demand functions: First order conditions for utility maximization in (2) lead to demand functions:

$$q_{it}=\varphi_{it}^{\sigma-1}\left(\frac{p_{it}}{P_t}\right)^{-\sigma_t}\frac{Y_t}{P_t} \quad \text{for} \quad i=1,2,...,N, \tag{5}$$

where $P_t=C\left(p_t,\varphi_t;\sigma\right)$ is the unit expenditure function in (3).

Expenditure share function: Let $w_{it}=p_{it}q_{it}\Big/\sum_{n=1}^{N}p_{nt}q_{nt}$ represent the budget share of *i-th* commodity in period *t*. Simple algebraic manipulation of (5) along with definition of expenditure share leads to:

$$w_{it} = \varphi_{it}^{\sigma-1}\left(\frac{p_{it}}{P_t}\right)^{1-\sigma}. \tag{6}$$

Equation (6) is a crucial relationship which allows us to back out values of demand shocks from observed price and quantity data, up to a factor of proportionality. Taking log of (6) and rewriting we have

$$\ln \varphi_{it} = \frac{\ln w_{it}}{(\sigma-1)} + \left(\ln p_{it} - \ln P_t\right). \tag{7}$$

Equation (6) leads to

$$\frac{\varphi_{it}}{\varphi_{1t}} = \frac{p_{it}}{p_{i1}}\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}. \tag{8}$$

Once elasticity of subsitution parameter, σ, is estimated, ratios of demand shocxks can be determined directly from observed prices and expenditur shares using (8) without further estimation. Normalization in (4) leads to full set of demand shocks.

2.3 The Sato-Vartia Index

The Sato (1976) and Vartia (1976) logarithmic price index is the first log-change index number that satisfied the factor reversal test[12]. Further Sato (1976) and Feenstra (1994) have shown that the S-V index is COLI for CES preferences. The S-V[13] index in its logarithmic form is given by

$$\ln SV_{st} = \sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{p_{it}}{p_{is}}\right) \quad where\ \ \omega^{*}_{ist} = \frac{w_{it}-w_{is}}{\ln w_{it} - \ln w_{is}} \Bigg/ \sum_{i=1}^{N} \frac{w_{it}-w_{is}}{\ln w_{it} - \ln w_{is}} \tag{9a}$$

The S-V index being a COLI for CES preferences, it is expected to be monotonically increasing in period *t* prices and also transitive. However Reinsdorf and Dorfman (1999) have shown that the S-V index fails monotonicity property and its lack of transitivity is shown in the numerical example in Section 3. We return to a discussion of these issues in Section 3.

Three other log-change index numbers are the Jevons (1863), geometric Young (1921) and the Tornqvist (1931) indices. These are respectively given by:

$$\ln JI_{st} = \frac{1}{N}\sum_{i=1}^{N}\left(\ln p_{it} - \ln p_{is}\right);\ \ \ln GYI_{st} = \sum_{i=1}^{N} w_i\left(\ln p_{it} - \ln p_{is}\right);\ \ln TI_{st} = \sum_{i=1}^{N}\left(\frac{w_{it}+w_{is}}{2}\right)\left(\ln p_{it} - \ln p_{is}\right) \tag{9b}$$

where $\left(w_i : i = 1, 2, ..., N\right)$ are a set of pre-specified weights which are constant for comparisons between all time periods.

---

[12] This means the product of price index numbers computed using S-V index number formula by just interchanging prices and quantities equals the change in expenditures. Abe and Rao (2019) have shown that the Sato-Vartia index is not unique.

[13] The Sato-Vartia index is COLI for CES preferences under constant taste parameters (Sato, 1976, Feenstra, 1994).

2.4 The Redding-Weinstein CES Unified Price Index

The R-W CES Unified Price Index (CUPI) is the center piece of this study. It is derived under the assumption of cardinal utility and heterogeneous CES preferences. Applying Konus (1924) concept of COLI for price comparisons from base period s to comparison period t [14], respectively, the R-W COLI is given by:

$$R-W\,COLI_{st}=\frac{E_t(p_t,U)}{E_s(p_s,U)}=\frac{C(p_t,\varphi_t,\sigma)}{C(p_s,\varphi_s,\sigma)}=\frac{\left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}}{\left(\sum_{i=1}^{N}\left(\frac{p_{is}}{\varphi_{is}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}} \tag{10}$$

The last equality follows from the unit expenditure function in (3).

We note in passing that COLI allowing for time-varying demand shocks can be derived for other utility and demand systems like the translog and AIDS and other invertible systems (see R-W, 2020, Section III) but we focus on CES preferences.
The main contribution of R-W (2020, p. 516) is to show that the COLI index in (10) can be written as the following logarithmic function of prices and quantities, labelled the *CES Unified Price Index (CUPI),*[15]:

$$\ln CUPI_{st}\left(p_t,p_s,q_t,q_s\right)=\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{p_{it}}{p_{is}}\right)-\left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right] \tag{11}$$

where

$$\omega_{ist}^{*}=\frac{w_{it}-w_{is}}{\ln w_{it}-\ln w_{is}}\Bigg/\sum_{i=1}^{N}\frac{w_{it}-w_{is}}{\ln w_{it}-\ln w_{is}} \tag{12}$$

and $\{\varphi_{i\tau}: i=1,2,...,N;\ \tau=s,t\}$ are determined using observed price and quantity data and equations (4) and (8).

The *CUPI* is linked to S-V index as (11) can be written as:

$$\ln CUPI_{st}=\ln SV_{st}-\left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right]. \tag{13}$$

We refer to the CUPI as generalized logarithmic index since it differs from other logarithmic indices shown in (9a) and (9b) due to the additional term in (11) measuring shifts in demand shock parameters in the two periods.

[14] This switch from two adjacent periods to any two periods allows us to introduce the notion of transitivity at a later stage.

[15] Here we deviate in notation from R-W and use $CUPI_{st}$ instead of Φ_{st}^{CUPI}.

3. Transitivity of Consumer Preferences, Duality and CUPI

In addition to the main R-W result that CUPI is COLI for CES with taste shocks, we establish additional economic theoretic properties including: transitivity which is required for transitivity of consumer preferences; and monotonicity of CUPI, a result consistent with COLI and duality theory.

3.1 Transitivity of CUPI and Preferences

Transitivity of consumer preferences is an important axiom in microeconomic theory which is critical to the existence of a utility function that preserves preference orderings over commodity bundles. Transitivity, referred to as circular test in Samuelson and Swamy (1974, p. 571), ensures consistency of price comparisons over different pairs of time periods. S-S use economic theoretic arguments to conclude that transitivity of price index numbers is essential for transitivity of real income or welfare comparisons.

The Konus (1924) index, ratio of expenditure functions under prices prevailing in two periods, say *t* and *s*,

$$COLI_{st}^{Konus} = \frac{E(p_t,U)}{E(p_s,U)} \tag{14}$$

is the theoretical basis for all cost of living index numbers. From (14) it is clear that this index satisfies transitivity, i.e.

$$COLI_{st}^{Konus} \times COLI_{tu}^{Konus} = COLI_{su}^{Konus} \quad for\ all\ s,t,u$$

From duality theory, it is known that the expenditure functions used in (14) are monotonically increasing in prices at *t*. Therefore, $COLI_{st}^{Konus}$ must be increasing in comparison period prices, p_t .

Consider three logarithmic price index numbers, the CUPI, the Tornqvist and the S-V indices. Diewert (1976) has shown that the Tornqvist is COLI for translog preferences; Sato (1976) and Feenstra (1994) have shown that the S-V index is COLI for CES preferences; and CUPI is COLI for CES with taste shocks (Redding and Weinstein, 2020). Consequently, we would expect all these three indices to satisfy transitivity and to be monotonically increasing in comparison period prices.

Reinsdorf and Dorfman (1999) have shown that the Tornqvist and S-V indices do not satisfy monotonicity globally. Numerical example in Table 1 below shows that both Tornqvist and S-V indices fail transitivity while CUPI satisfies transitivity.

Table 1: Lack of Transitivity of Tornqvist and SV indices

Data				Price Indices			
Time	Commodiy	Price	Quantity	Time Pair	S-V	Tornqvist	CUPI
1	1	1000	2	P(1,2)	0.3476	0.3409	0.3086
	2	200	20	P(2,3)	3.7971	3.6711	3.9173
	3	150	30	P(1,3)	1.4322	1.4058	1.2089
2	1	500	3	P(1,2)×P(2,3)	1.3198	1.2515	1.2089
	2	100	30				
	3	20	40				
3	1	2000	10				
	2	100	60				
	3	300	70				

The failure of S-V and Tornqvist indices to satisfy transitivity, illustrated through this example, is a puzzle – the puzzle being that as COLI both of these must satisfy transitivity but when applied to observed price and quantity data these are no longer transitive. However CUPI satisfies transitivity consistent with it being COLI. In what follows below we formally prove that CUPI satisfies transitivity. Explanation of the puzzle lies in the fact that although COLI in (14) does not directly depend on observed quantity vectors, the observed quantities are expected to be on the demand function implied by the utility function. However observed quantities are likely to fail this due to a variety of reasons including: time-varying taste shocks; quality differences; stockpiling behaviors, habit formation, and pure random shocks reflecting inherent stochastic behavior.

For a formal explanation, consider the case of identical taste parameters in both periods, $\varphi_{is} = \varphi_{it} = \varphi_i$, assumption underlying the S-V index from CES preferences. Equation (8) implies, for any two commodities i and j and for periods s and t:

$$\begin{aligned}\ln\varphi_i - \ln\varphi_j &= \ln p_{it} - \ln p_{jt} - \frac{1}{1-\sigma}\left(\ln w_{jt} - \ln w_{it}\right) \\ &= \ln p_{is} - \ln p_{js} - \frac{1}{1-\sigma}\left(\ln w_{js} - \ln w_{is}\right)\end{aligned} \tag{15}$$

This is an exact relation between observed prices and expenditure shares. If observed quantity and price data do not satisfy (15), the computed S-V index differs from COLI for CES preferences, hence the failure of S-V index to pass transitivity test.

If we allow for changes in demand shocks over time, then it is possible to find a vector of demand shocks that ensures consistency between observed prices and quantities and the first order conditions for consumer's utility maximization. This is the reason why CUPI is transitive. We state and prove this result formally.

Result 1: For any t, s and u, the R-W CUPI index with a given elasticity of substitution, σ *, satisfies:*

$$CUPI_{st}(p_s, q_s, p_t, q_t) \times CUPI_{tu}(p_t, q_t, p_u, q_u) = CUPI_{su}(p_s, q_s, p_u, q_u) \tag{16}$$

Proof: From the properties of CES function with time varying demand shocks and from equation (7), we have

$$\ln P_t = -\ln\varphi_{it} + \ln p_{it} - \frac{1}{1-\sigma}\ln w_{it}\,; \; i = 1, 2, ..., N$$

Then, the following relationship between $CUPI_{st}$ and P_s and P_t follows.

$$\begin{aligned}\ln P_t - \ln P_s &= \sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln P_t - \ln P_s\right) \\ &= \sum_{i=1}^{N}\omega_{ist}^{*}\left(-\ln\varphi_{it} + \ln p_{it} - \frac{1}{1-\sigma}\ln w_{it}\right) - \sum_{i=1}^{N}\omega_{ist}^{*}\left(-\ln\varphi_{is} + \ln p_{is} - \frac{1}{1-\sigma}\ln w_{is}\right) \\ &= \sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln p_{it} - \ln p_{is}\right) - \sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln\varphi_{it} - \ln\varphi_{is}\right) = CUPI_{st}\end{aligned} \tag{17}$$

From equation (17), $CUPI_{st} = P_t / P_s$ satisfies transitivity condition in equation (16).

This particular property of CUPI makes it analytically superior to the current gold standard, the Fisher's *ideal index* as the Fisher index does not satisfy transitivity. In view of transitivity and additional properties established below, the CUPI may be considered a theoretically complete index.

3.2 Monotonicity of COLI and CUPI

From duality theory, COLI must be monotonically increasing in comparison period prices. Monotonicity states that if period *t* price of one of the commodities increases, with all other prices remaining the same, the price index must record an increase. Result 2 establishes monotonicity of CUPI.

Result 2: Suppose price of commodity i at time t increases by $\varepsilon\ (>0)$ *with all other prices remaining the same, with the new price vector* p_t' *:*

$$p_t' = \left(p_{1t}, p_{2t}, .., p_{it} + \varepsilon, p_{i+1t}, .., p_{Nt} \right) \quad \textit{with} \ \ \varepsilon > 0$$

then

$$CUPI_{st}(p_s, q_s, p_t, q_t) \ < \ CUPI_{st}(p_s, q_s, p_t + \varepsilon, q_t)$$

Proof: To prove this, consider equation (17) above

$$\ln CUPI_{st} = \ln P_t - \ln P_s$$

where $\ln P_t = -\ln \varphi_{it} + \ln p_{it} - \dfrac{1}{1-\sigma} \ln w_{it}$ and $\ln P_s = -\ln \varphi_{is} + \ln p_{is} - \dfrac{1}{1-\sigma} \ln w_{is}$. We note that P_s is independent of prices in period *t*. Because both w_{it} and φ_{it} are differentiable functions with respect to p_{it} . .To establish monotonicity, it is sufficient to show that:

$$\frac{d \ln P_t}{d \ln p_{it}} = -\frac{d \ln \varphi_{it}}{d \ln p_{it}} + 1 - \frac{1}{1-\sigma} \frac{d \ln w_{it}}{d \ln p_{it}} \ > 0. \tag{18}$$

It is tedious but straightforward to find derivates in (18) to establish this result. Detailed derivations are in Appendix A1.

In addition to these important theoretical properties of monotonicity and transitivity, CUPI also possesses additional properties expected of aggregate measures of price change such as linear homogeneity, identity and it satisfies factor reversal test.

Linear homogeneity states that if prices in period *t* are all multiplied by a positive scalar then the measure of price change itself is multiplied by the same scalar.

Result 3: For any $\lambda > 0$, $CUPI_{st}(p_s, q_s, \lambda p_t, q_t) = \lambda \cdot CUPI_{st}(p_s, q_s, p_t, q_t)$

Proof: Recall that $CUPI_{st}$ in logarithmic form is

$$\ln CUPI_{st} = \sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{p_{it}}{p_{is}}\right) - \left[\sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right] = \sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{p_{it}/\varphi_{it}}{p_{is}/\varphi_{is}}\right)$$

Suppose the taste parameters associated with new prices in period *t*, λp_t, are denoted by $\{\varphi^{*}_{it} : i = 1,2,...,N\}$ which means that $CUPI(p_s, q_s, \lambda p_t, q_t)$ can be written as:

$$\ln CUPI_{st} = \sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{\lambda p_{it}}{p_{is}}\right) - \left[\sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{\varphi^{*}_{it}}{\varphi_{is}}\right)\right] = \sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{p_{it}/\varphi^{*}_{it}}{p_{is}/\varphi_{is}}\right)$$

It is sufficient if we show that $\{\varphi_{it} = \varphi^{*}_{it} : i = 1,2,...,N\}$. Using equation (8) and the geometric normalization, it is easy to show that φ_{it} and φ^{*}_{it} for $i = 1,2,...,N$ are, respectively, given by:

$$\varphi_{it} = \left(\frac{p_{it}}{p_{1t}}\right)\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}} \left[\prod_{k=2}^{N} \left\{\left(\frac{p_{kt}}{p_{1t}}\right)\left(\frac{w_{kt}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\right\}^{(-1/N)}\right]\varphi \tag{19}$$

and

$$\varphi^{*}_{it} = \left(\frac{\lambda p_{it}}{\lambda p_{1t}}\right)\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}} \left[\prod_{k=2}^{N} \left\{\left(\frac{\lambda p_{kt}}{\lambda p_{1t}}\right)\left(\frac{w_{kt}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\right\}^{(-1/N)}\right]\varphi$$

which establishes equality of taste parameters φ_{it} and φ^{*}_{it}.

Now we turn to *identity* property which requires the aggregate measure of price change to equal 1 when prices and quantities are the same in both periods, i.e., $\{p_{it} = p_{is}; q_{it} = q_{is} : i = 1,2,...,N\}$. This property is sometimes called weak identity test[16].

Result 4: The $CUPI_{st}$ in equations (11) satisfies identity test when prices and quantities remain unchanged over time.

Proof: When prices and quantities are the same in both periods, from (19) we have $\{\varphi_{it} = \varphi_{is} \ \ for \ i = 1,2,...,N\}$. This implies

$$\ln CUPI_{st} = \sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{p_{it}}{p_{is}}\right) - \left[\sum_{i=1}^{N} \omega^{*}_{ist} \ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right] = 0 \quad when \ \ p_{it} = p_{is} \ \ for \ i = 1,2,...,N \Rightarrow CUPI_{st} = 1.$$

[16] Strong identity test requires the price index to equal 1 when prices in both periods are the same but allowing for different quantity vectors.

3.4 Implicit CUPI Quantity Index and Welfare Comparisons

We start with a quote from Samuelson and Swamy (1974, p. 567), "Although most attention in the literature is devoted to price indexes, when you analyze the use to which price indexes are generally put, you realize that quantity indexes are actually most important. Once somehow estimated, price indexes are in fact used, if at all, primarily to "deflate" nominal or monetary totals in order to arrive at estimates of underlying "real magnitudes" (which is to say, quantity indexes!)."

Implicit measures of quantity change can be obtained by invoking the fundamental decomposition of value change (ratio of expenditures in the two periods) into price and quantity changes. We compute indirect quantity comparisons using $CUPI_{st}$ as

$$Q_{st}^{\text{CUPI}} = \frac{\sum_{i=1}^{N} p_{it} q_{it} \Big/ \sum_{i=1}^{N} p_{is} q_{is}}{CUPI_{st}} = \frac{\sum_{i=1}^{N} p_{it} q_{it} \Big/ CUPI_{st}}{\sum_{i=1}^{N} p_{is} q_{is}} \tag{20}$$

After simple algebraic manipulation, the implicit quantity index in (20), in logs, is given by:

$$\ln Q_{st}^{CUPI} = \sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{q_{it}}{q_{is}}\right) + \left[\sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right] \tag{21}$$

The CUPI quantity index in (21) is similar in structure to the CUPI in (5). Further, by construction, product of these two indices equals the ratio of expenditures in the two periods. Rewriting the expressions for CUPI price and quantity indices, we can show that these indices satisfy the *factor reversal test*. The CUPI price and quantity indices (5) and (21) can be rewritten as:

$$\begin{aligned} \ln CUPI_{st} &= \sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{p_{it}}{p_{is}}\right) - \left[\sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right] = \sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{p_{it}/\varphi_{it}}{p_{is}/\varphi_{is}}\right) \\ &= \sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{p_{it}^{*}}{p_{is}^{*}}\right) \quad \text{where } p_{i\tau}^{*} = p_{i\tau}/\varphi_{i\tau} \quad \tau = t, s \end{aligned}$$

and

$$\begin{aligned} \ln Q_{st}^{CUPI} &= \sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{q_{it}}{q_{is}}\right) + \left[\sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right] \\ &= \sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{q_{it}\varphi_{it}}{q_{is}\varphi_{is}}\right) = \sum_{i=1}^{N} \omega_{ist}^{*} \ln\left(\frac{q_{it}^{*}}{q_{is}^{*}}\right) \quad \text{where } q_{i\tau}^{*} = q_{i\tau}\varphi_{i\tau} \quad \tau = s, t \end{aligned}$$

From these two equations it can be seen that the CUPI quantity index can be simply obtained by replacing ratios of transformed prices $\left\{p_{i\tau}^{*} : i = 1, 2, ..., N;\ \tau = s, t\right\}$ with ratios of transformed quantities $\left\{q_{i\tau}^{*} : i = 1, 2, ..., N;\ \tau = s, t\right\}$.

The following result provides a welfare comparison interpretation to the implicit CUPI quantity index.

Result 5: Under cardinal utility framework and homothetic CES preferences with demand shocks and under the assumption that the observed quantities in periods *t* and *s* are cost-minimizing under the prevailing prices, p_t and p_s, and utility levels, U_t and U_s, the CUPI quantity index in (20) equals the ratio of utility levels in these two periods thus providing a measure of welfare change.

Proof: Since preferences are CES and homothetic and under the assumption of optimality of observed quantities, we have

$$\sum_{i=1}^{N} p_{i\tau} q_{i\tau} = E_{\tau}(p_{\tau}, U_{\tau}) = \left(\sum_{i=1}^{N} (a_{i\tau})^{\sigma} (p_{i\tau})^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \times U_{\tau} \qquad \tau = s,t \tag{22}$$

From (22) and the definition of COLI and CUPI price index, we have

$$\ln \sum_{i=1}^{N} p_{it} q_{it} - \ln \sum_{i=1}^{N} p_{is} q_{is} = \ln CUPI_{st} + \ln U_t - \ln U_s \tag{23}$$

Since the CUPI price index (11) and CUPI quantity index (21) satisfy factor reversal test we have:

$$\ln \sum_{i=1}^{N} p_{it} q_{it} - \ln \sum_{i=1}^{N} p_{is} q_{is} = \ln CUPI_{st} + \ln Q_{st}^{CUPI} \tag{24}$$

Equations (23) and (24) together imply:

$$\ln Q_{st}^{CUPI} = \ln U_t - \ln U_s \;\Rightarrow\; Q_{st}^{CUPI} = \frac{U_t}{U_s} = \textit{welfare change}$$

This result shows that the implicit quantity index derived by deflating change in expenditures by CUPI provides a measure of welfare change.

3.5 CUPI: A solution to the practical Chain Drift problem with Scanner Data

Recently, more and more researchers and statistical offices in various countries are incorporating high frequency point of sale scanner data into compilation of aggregate measures of price change. Often they encounter the problem of chain drift when chained price index numbers are used. Diewert (2020, p.3) describes: "…*Chain drift* occurs when an index does not return to unity when prices in the current period return to their levels in the base period". The seriousness of chain drift problem is well documented in Diewert (2018, 2020)[17]. The spike in purchases and then return to the original levels is a major driver of the chain drift problem. Feenstra and Shapiro (2003) considered the chain drift problem caused by sales and stocking behaviour and suggested the use of fixed base indices as a solution. However fixed base index numbers have problem including sensitivity of price change measures to the choice of the base period.

COLI like the Fisher, Tornqvist and S-V indices exhibit significant chain drift when using scanner data (de Haan and van der Grient, 2011). To date, solutions to this problem rely heavily on multilateral index number formulae which satisfy transitivity. The use of Gini-Elteto-Koves-Szulc (GEKS) (Gini, 1931; Elteto and Koves, 1964; Szulc, 1964) index and the rolling-window GEKS methods (Ivancic, Diewert and Fox, 2009 and 2011; de Haan, 2008; de Haan and Krsinich, 2014; van Auer, 2019) and variants of the Geary (1958) and Khamis (1972) method have been proposed and being explored (Chessa, 2016; Lamboray, 2017). For example, the GEKS method builds on the Fisher binary index (Diewert, 1976) but has limited economic theoretic interpretation (Neary, 2004). The GEKS method at best can be considered as a technique that generates transitive comparisons which deviate the least from a set of non-transitive bilateral comparisons.

[17] Diewert (2020) will be included as Chapter 7 in the latest version of the *CPI Manual* (ILO/IMF/OECD/UNECE/Eurostat/The World Bank, 2020).

These methods offer viable solutions to the problem lack but solid economic theoretic foundations and remain essentially heuristic.

The CUPI, in addition to being a COLI and transitive, has an additional practical advantage over the alternatives when it comes to transitive multilateral comparisons. The multilateral methods, listed in the previous paragraph, have a serious practical disadvantage that price comparisons need to be completely revised whenever a new time period is added. This is a major problem for economic statisticians measuring aggregate price changes. The CUPI (in equation 11), in contrast, not only satisfies transitivity but CUPI comparisons between any two periods are functions of data for only those two periods and not affected by data from other periods hence does not need revisions when new data are added.

In light of this discussion, our result that CUPI is a transitive index assumes considerable practical significance. Thus CUPI meets analytical requirements outlined in Samuelson and Swamy (1974) and at the same time serves as a solution to the practical problem of chain drift discussed in Diewert (2020). In contrast, the S-V index, the Fisher and Tornqvist indices exhibit severe chain drift – a point we return to in our empirical results section.

4. R-W Geometric Normalization and Commensurab ility

We now return to the problem of choice of normalization rule to work with CES utility function with demand shocks. R-W (2020, Section II F) propose geometric normalization but also discuss alternative normalizations. In this section we prove that their geometric normalization belongs to the class of admissible normalizations if the resulting price change measures need to be independent of units of measurement or satisfy commensurability. Price indices must be independent of measurement units of commodities. Otherwise, choice of a measurement unit such as litre or gallon, kilogram or pound, affects the index number value. The following numerical example shows that CUPI with geometric normalization rule in (4) satisfies this property whereas CUPI index with the additive normalization rule, $\sum_{i=1}^{N} \varphi_{it} \Big/ N = \varphi$ does not.

Table 2: Measurement Units and the CUPI Index – An illustrative example

Panel A

Data				Price Indices			
Time	Commodiy	Price	Quantity	Time Pair	S-V	CUPI (1)	CUPI (2)
1	1	1000	2	P(1,2)	0.3476	0.3086	0.4753
	2	200	20	P(2,3)	3.7971	3.9173	4.2090
	3	150	30	P(1,3)	1.4322	1.2089	2.0005
2	1	500	3				
	2	100	30				
	3	20	40				
3	1	2000	10				
	2	100	60				
	3	300	70				

Panel B

Data				Price Indices			
Time	Commodiy	Price	Quantity	Time Pair	S-V	CUPI (1)	CUPI (2)
1	1	100	20	P(1,2)	0.3476	0.3086	0.3796
	2	200	20	P(2,3)	3.7971	3.9173	3.5122
	3	150	30	P(1,3)	1.4322	1.2089	1.3333
2	1	50	30				
	2	100	30				
	3	20	40				
3	1	200	100				
	2	100	60				
	3	300	70				

Note: σ is set to be 3.
CUPI (1) uses the unweigted geometric means as the normalization condition
CUPI (2) uses the unweighted arithmetice means as the normalization condition
P(1,2) is the price index from time 1 to 2.

In this example, Panels A and B represent the same price data except that in Panel B the unit of measurement of the first commodity is changed. The CUPI based on geometric normalization, CUPI(1), is the same in both panels. However, CUPI (2) with additive normalization leads to an index value of 0.3796 compared to 0.4753 before change in units of measurement – a difference of 20 percent! This illustrates that CUPI with arithmetic normalization rule is not independent of units of measurement.

Result 6: The CUPI with geometric normalization of demand shock parameters is independent of units of measurement.

Proof: For any $p_t, p_s, q_t, q_s \in \mathbf{R}^N_{++}$, consider a positive valued vector $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_N\} \in \mathbf{R}^N_{++}$. Then price and quantity vectors after change in units of measurement are:

$$p_t^* = (\lambda_1 p_{1t}, \lambda_2 p_{2t}, ..., \lambda_N p_{Nt}), \qquad p_s^* = (\lambda_1 p_{1s}, \lambda_2 p_{2s}, ..., \lambda_N p_{Ns}), \qquad q_t^* = \left(\frac{q_{1t}}{\lambda_1}, \frac{q_{2t}}{\lambda_2}, ..., \frac{q_{Nt}}{\lambda_N}\right),$$

$q_s^* = \left(\frac{q_{1s}}{\lambda_1}, \frac{q_{2s}}{\lambda_2}, ..., \frac{q_{Ns}}{\lambda_N}\right)$ and new demand shock parameters $(\varphi_{1s}^*, \varphi_{2s}^*, \ldots \varphi_{Ns}^*)$ and $(\varphi_{1t}^*, \varphi_{2t}^*, \ldots \varphi_{Nt}^*)$

Expenditure shares remain unchanged after change in units of measurement. The CUPI before and after change in units of measurement:

$$\ln CUPI_{st}\left(p_s,\, q_s,\, p_t,\, q_t\right) = \left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{p_{it}}{p_{is}}\right)\right] - \left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right]$$

$$\ln CUPI_{st}\left(p_s^{*},\, q_s^{*},\, p_t^{*},\, q_t^{*}\right) = \left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{p_{it}^{*}}{p_{is}^{*}}\right)\right] - \left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{\varphi_{it}^{*}}{\varphi_{is}^{*}}\right)\right]$$

Since the first expression on the right hand side is independent of units of measurement, it is sufficient if we show that

$$\left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{\varphi_{it}}{\varphi_{is}}\right)\right] = \left[\sum_{i=1}^{N}\omega_{ist}^{*}\ln\left(\frac{\varphi_{it}^{*}}{\varphi_{is}^{*}}\right)\right] \Leftrightarrow \left[\sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln\varphi_{it} - \ln\varphi_{is}\right)\right] = \left[\sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln\varphi_{it}^{*} - \ln\varphi_{is}^{*}\right)\right]$$

Or equivalently,

$$\left[\sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln\varphi_{it} - \ln\varphi_{it}^{*}\right)\right] = \left[\sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln\varphi_{is} - \ln\varphi_{is}^{*}\right)\right] \tag{25}$$

Taking logarithms equation (8) to data before and after change in units of measurement we have for period t and i=1,2,…,N

$$\ln\varphi_{it} = \ln p_{it} - \ln p_{1t} + \left(\frac{1}{\sigma - 1}\right)\left(\ln w_{it} - \ln w_{1t}\right) + \ln\varphi_{1t} + \ln\varphi. \tag{26a}$$

$$\begin{aligned}\ln\varphi_{it}^{*} &= \ln p_{it}^{*} - \ln p_{1t}^{*} + \left(\frac{1}{\sigma - 1}\right)\left(\ln w_{it} - \ln w_{1t}\right) + \ln\varphi_{1t}^{*} + \ln\varphi \\ &= \ln p_{it} + \ln\lambda_i - \ln p_{1t} - \ln\lambda_1 + \left(\frac{1}{\sigma - 1}\right)\left(\ln w_{it} - \ln w_{1t}\right) + \ln\varphi_{1t}^{*} + \ln\varphi.\end{aligned} \tag{26b}$$

Given (26a) and (26b), to establish (25) it is sufficient if we show that:

$$\ln\varphi_{1s} - \ln\varphi_{1s}^{*} = \ln\varphi_{1t} - \ln\varphi_{1t}^{*}. \tag{27}$$

From the geometric normalization used for CUPI and equation (8), we have

$$\varphi_{1t} = \left[\prod_{k=2}\left(\frac{p_{kt}}{p_{1t}}\right)\left(\frac{w_{kt}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\right]^{-1/N}\varphi \;\; and \;\; \varphi_{1t}^{*} = \left[\prod_{k=2}^{N}\left(\frac{p_{kt}}{p_{1t}}\right)\left(\frac{w_{kt}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\left(\frac{\lambda_k}{\lambda_1}\right)\right]^{-1/N}\varphi = \varphi_{1t}\left[\prod_{k=2}^{N}\left(\frac{\lambda_i}{\lambda_1}\right)\right]^{1/N}\varphi \tag{28}$$

Equation (27) follows from (28) and hence the result.

Commensurability and CUPI Geometric normalization

Here we provide a necessary and sufficient condition on the normalization on shock parameters for CUPI to be independent from measurement units. Equation (8) shows that the shock parameter for commodity *i* in period $\tau\ (=s,t)$ is related to prices and expenditure shares through:

$$\varphi_{i\tau} = \left(\frac{p_{i\tau}}{p_{1\tau}}\right)\left(\frac{w_{i\tau}}{w_{1\tau}}\right)^{\frac{1}{\sigma-1}} \varphi_{1\tau} \quad i=1,2,...,N;\ \tau=s,t \tag{29}$$

Any normalization can be expressed in the following general form:

$$\varphi_{1\tau} = f\left(\varphi_{2\tau}, \varphi_{3\tau}, \ldots, \varphi_{N\tau}\right) \quad \tau=s,t \tag{30}$$

Substituting (29) into (30) leads to

$$\varphi_{1\tau} = f\left(x_{2\tau}\varphi_{1\tau}, x_{3\tau}\varphi_{1\tau}, \ldots, x_{N\tau}\varphi_{1\tau}\right) \quad where\ \ x_{i\tau} = \left(\frac{p_{i\tau}}{p_{1\tau}}\right)\left(\frac{w_{i\tau}}{w_{1\tau}}\right)^{\frac{1}{\sigma-1}} \ for\ all\ i=1,2,...,N\ and\ \tau=s,t \tag{31}$$

Let solution for $\varphi_{1\tau}\ (\tau=s,t)$ in (31) be represented as:

$$\varphi_{1\tau} = g\left(x_{2\tau}, x_{3\tau}, \ldots, x_{N\tau}\right) \quad \tau=s,t \tag{32}$$

where $g:\Re_{++}^{N-1} \to \Re_{++}$ is a positive real-valued function.

Result 7: The CUPI index defined in (10) and (11) is independent from tmeasurement units *if and only if* the normalization condition can be expresseds in the form

$$\varphi_1 = A \times \left(\prod_{i=2}^{N} x_{it}^{c_i}\right) \text{ where } c_i \in \mathbf{R},\ A>0 \tag{33}$$

If all the taste parameters, φ_i, are considered equally important and treated symmetrically, the necessary and sufficient condition can be written in the form of a simple geometric mean, which is identical to the normalization condition by R-W (2020),

$$\prod_{i=1}^{N}(\varphi_i) = \varphi \tag{34}$$

where φ is a positive constant.

Proof: This result is proved using mathematical induction. We prove the result for *N = 2*. Assuming that this is true for a general *N* and we prove that the result holds for *N*+1. Proof for *N*=2 is presented here and rest of the proof is in Appendix 2.

Eaquation (8) applied to N = 2 gives us:

$$\varphi_{2t} = \left(\frac{p_{2t}}{p_{1t}}\right)\left(\frac{w_{2t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}} \varphi_{1t}.$$

When change in measurement units occur, the new price, p_{it}^*, and quantity, q_{it}^*, are,

$$p_{it}^* = p_{it}\lambda_i,\ \ q_{it}^* = q_{it}/\lambda_i \quad \text{where } \lambda_i > 0 \text{ for } i = 1 \text{ or } 2.$$

Let φ_{1t}^* denote the taste parameter for commodity 1 after change in measurement units. Then from equation (27), the necessary and sufficient condition for commensurability is,

$$\ln\varphi_{1s} - \ln\varphi_{1s}^* = \ln\varphi_{1t} - \ln\varphi_{1t}^*.$$

Let $\ln K$ denote the difference,

$$\ln\varphi_{1s} - \ln\varphi_{1s}^* = \ln\varphi_{1t} - \ln\varphi_{1t}^* = \ln K \tag{35}$$

Following (31), denote

$$x_t = \left(\frac{p_{2t}}{p_{1t}}\right)\left(\frac{w_{2t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}};\ x_t^* = \left(\frac{p_{2t}^*}{p_{1t}^*}\right)\left(\frac{w_{2t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}} \ \text{ and } \ \lambda = \left(\frac{\lambda_2}{\lambda_1}\right).$$

The normalization conditions before and after change in units of measurement are:

$$\varphi_{1t} = g\left(x_t\right);\ \ \varphi_{1t}^* = g\left(x_t^*\right) = g\left(\lambda x_t\right). \tag{36}$$

Using (36), equation (35) can be written as

$$\ln g\left(\lambda x_t\right) - \ln g\left(x_t\right) = \ln g\left(\lambda x_s\right) - \ln g\left(x_s\right) = \ln m. \tag{37}$$

Equation (37) must hold for all values of x_t and x_s. Therefore, m depends on λ but not on x_t or x_s.

Therefore, If the index passes the commensurability test, given λ, there is a constant $m(\lambda) > 0$ that satisfies the following equation for all x_t

$$g\left(x_t\right) = \frac{1}{m(\lambda)} g\left(\lambda x_t\right) \ \Rightarrow g\left(\lambda x_t\right) = g(x_t)m(\lambda)$$

Letting $f(\lambda x_t) = g(\lambda x_t)$, we have

$$f(\lambda x_t) = g(x_t)\cdot m(\lambda) \tag{38}$$

Equation (38) is a well-known functional equation whose general non-zero solution for $x_t > 0$ and $\lambda > 0$ is given by[18]

$$f\left(x_t\right) = abx_t^c;\ \ g\left(x_t\right) = ax_t^c,;\ \ m\left(\lambda\right) = b\lambda^c,\ \ a,b,c \in \mathbf{R}\ .$$

Therefore, functional form for $g\left(x\right)$ combined with equation (8) yields

[18] The derivation of the general solution for this functional equation can be found in Aczel (1966).

$$\varphi_{1t} = a x_t^c \quad where \quad x_t = \left(\frac{p_{2t}}{p_{1t}}\right)\left(\frac{w_{2t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}} \tag{39}$$

This establishes the result for $N = 2$. Rest of the proof based on induction is presented in Appendix A2.

Corollary: The geometric normalization used in R-W specification satisfies the condition stated in Result 7. Additive normalization discussed in R-W (2020, p. 518) does not satisfy this condition and hence fails commensurability condition.

Proof: Geometric mean normalization implies, for each, $\tau (= s,t)$

$$\prod_{i=1}^{N}(\varphi_{i\tau})^{1/N} = \varphi \Rightarrow \varphi_{1\tau} = \varphi \prod_{i=2}^{N}(x_{i\tau})^{-1/N} \quad for \ \tau = s,t \tag{40}$$

This satisfies the condition in equation (39) and hence the index is independent of units of measurement.

In the case of arithmetic mean normalization, we have for $\tau = s,t$

$$\frac{\sum_{i=1}^{N}\varphi_{i\tau}}{N} = 1 \ \Rightarrow \ \varphi_{1\tau} = N - \sum_{i=2}^{N}\varphi_{i\tau} \ \Rightarrow \ \varphi_{1\tau} = N - \varphi_{1\tau}\sum_{i=2}^{N}x_{i\tau} \ \Rightarrow \varphi_{1\tau} = \frac{N}{1+\sum_{i=2}^{N}x_{i\tau}} \tag{41}$$

The last expression on the right side in (41) cannot be written in the form (39) and hence the CUPI with arithmetic mean normalization fails commensurability condition. This is also true for other generalized specifications discussed in R-W (2020, p. 518).

The main conclusion of this section is that CUPI with geometric normalization is unique in that CUPI with any other normalization would violate invariance with respect to units of measurement.

5. Empirical Results: CUPI and Scanner Data from Japan

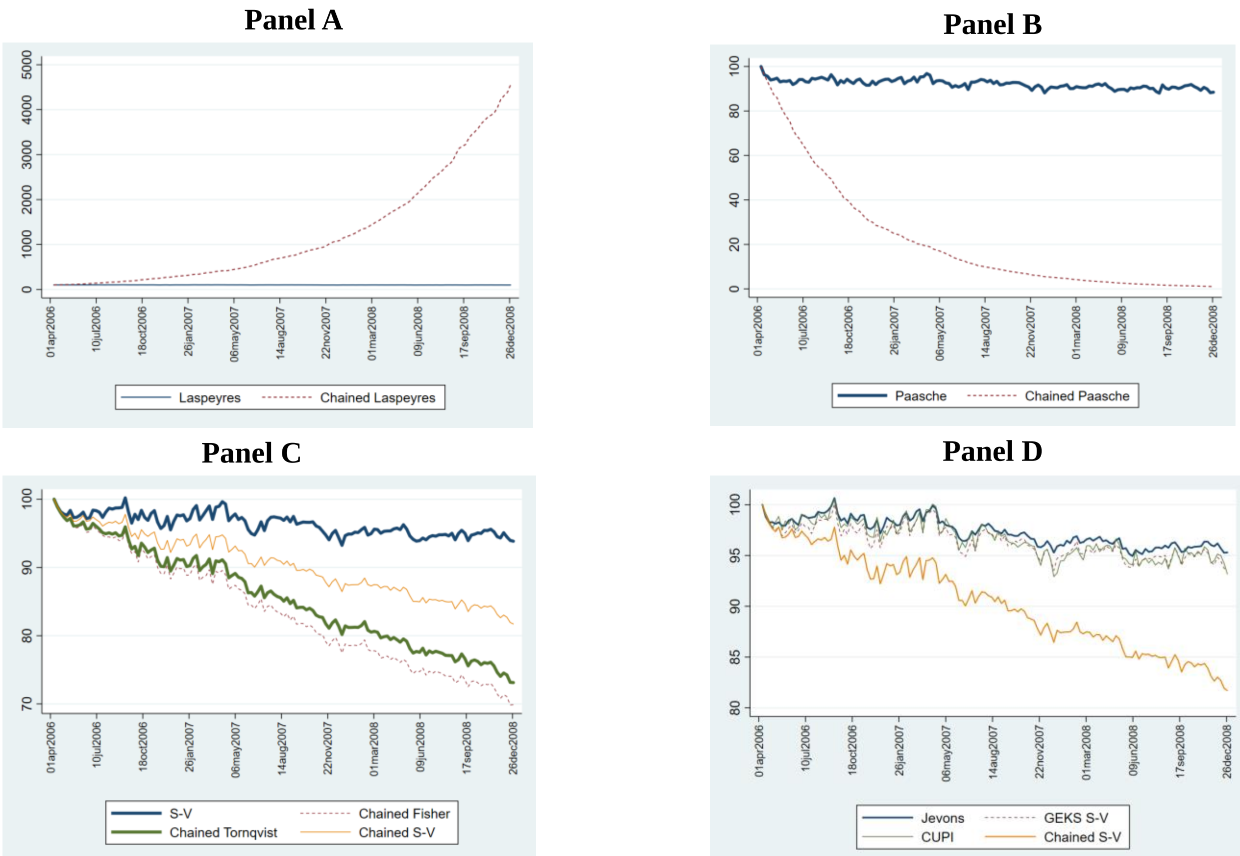
We use scanner data from nationwide retails stores in Japan provided by the Intage Holdings Inc.[19] The data set contains barcode level weekly sales and quantity data from retail stores all over Japan. For purposes of illustration, we have opted to use weekly data from April 2006 and December 2008, which gives us 143 weeks of observations[20]. As our main objective is to examine the nature and extent of chain drift and the performance of CUPI index against superlative indices, we have decided to make use of weekly data. In this paper, we use scanner data on 41 different cereal items identified by Japanese Article Number (JAN) that were sold in 428 retail stores throughout the study period. Items sold in different stores are treated as separate commodities even though they have identical barcodes, resulting in 1,012 pairs of commodity and stores each week[21], and 144,716 observations.

[19] The Research Center for Economic and Social Risks at Hitotsubashi University houses scanner data collected and provided by Intage Holdings Inc. For further information on the data set available, see hiips://www.intage.co.jp/english/.

[20] We opted for this period since Intage has allowed us to share this data with other users and readers.

[21] Not all items are sold in all the outlets.

Figure 1: Measures of Price Change using the Laspeyres, Paasche, Fisher, Sato-Vartia indexes, Chained indices and the CES Unified Price Index
(base period: first week, April 2009)



Note: Numerical values of GEKS S-V, GEKS Fisher and GEKS-Tornqvist are very close and therefore only GEKS S-V is presented in Panel D.

Price indices in these charts show price changes with first week (April, 2009) as the base period. Panels A and B in Figure 1 show the chain drift associated with the Laspeyres and Paasche indices, the drifts moving in opposite directions. Panel C shows price changes measured using fixed-base S-V index and chained Fisher, Tornqvist and S-V indices. Chained indices in Panel C show a decline in prices around 30 percent compared to a decline of only 4 percent when fixed base S-V index is used. This magnitude of large chain drift associated with superlative indices and the S-V index is consistent with results reported in de Haan and van der Grient (2011) based on weekly scanner data in Netherlands. Similar magnitudes of chain drift are also reported in the numerical examples presented in Diewert (2020, Appendix).

Conventional approach to tackle chain drift problem is to use index number formulae that are transitive. In this section we consider the following methods: three variants of GEKS method using the Fisher, Tornqvist and S-V binary indices; the Jevons index which is an unweighted geometric average of price changes, and finally CUPI which is a transitive index by construction. The elasticity of substitution, σ, needed for CUPI computation is estimated using Feenstra (1994) algorithm and found to be 7.0894. Our estimate is between the 50^{th} and 75^{th} percentile of estimates reported by R-W (2020, p. 534). A comparison of price change measures with and without accounting for time varying demand shocks, we find that the S-V index exhibits a bias of 0.2233 per cent per annum compared to CUPI. Compared with CUPI, the conventional chained superlative indices have a negative bias of the order 4.875 per cent per annum. Further details of differences are in Table A1 in Appendix A3.

In summary, the CUPI index performs well numerically compared to measures of aggregate price change computed using other methods currently in vogue. CUPI is superior to these methods due its theoretical properties and the fact that CUPI based aggregate price measures need not be revised when new data points are added is a distinct advantage of CUPI.

6. Conclusions

The primary objective of the paper is to present a method of aggregating price changes that is theoretically sound and yet practical. We have shown that the generalized logarithm index with demand shocks represented by the CES unified price index (CUPI), proposed by Redding and Weinstein (2020), is a method that meets both analytical and practical requirements. Our most significant finding is that the CUPI satisfies monotonicity and also transitivity of price comparisons, a condition Samuelson and Swamy (1974) considered to be essential for transitivity of consumer preferences. Thus CUPI meets an important theoretical requirement which is not met by traditional indices like the Fisher, Tornqvist and S-V indices. As a corollary, CUPI offers a practical solution for applications with scanner data as it eliminates chain drift. Another significant feature of CUPI is that price comparisons are not affected when new data are introduced – this is not the case with currently used transitive methods. We also show that CUPI with geometric normalization of taste parameters, which is the preferred option of R-W (2020), belongs to the class of normalizations that to an index that is independent of units of measurement. We are also able to establish additional properties of CUPI, such as linear homogeneity and the factor reversal test which underscore the versatility of CUPI. The CUPI with all these properties is vastly superior to Fisher, Tornqvist, S-V and other methods and it goes a long way in bridging the gap between index number theory and practice.

References

Abe, N. and D.S. P. Rao (2019) "Multilateral Sato–Vartia index for international comparisons of prices and real expenditures," Economics Letters, Volume 183,108535.

Chessa, A. (2016), "A new methodology for processing data in the Dutch CPI", Eurona, No. 1.

Crawford, Ian; Neary, J. Peter (2019), New characteristics and hedonic price index numbers, CESifo Working Paper, No. 7529, Center for Economic Studies and Ifo Institute (CESifo), Munich

Demnuynck, Thomas and Mathie Parenti (2018), "A price index with variable mark-ups and changing variety" NBER Working Paper No. 356.

de Haan, J. (2008), "Reducing Drift in Chained Superlative Price Indexes for Highly Disaggregated Data", paper presented at the Economic Measurement Workshop, Centre for Applied Economic Research, University of New South Wales, December 10.

de Haan, Jan, and Heymerik A. van der Grient (2011), "Eliminating chain drift in price indexes based on scanner data" Journal of Econometrics, Volume 161, Issue 1, Pages 36-46.

de Haan, J. and F. Krsinich (2014), Scanner Data and the Treatment of Quality Change in Nonrevisable Price Indexes, Journal of Business & Economic Statistics, 32:3, 341-358.

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

Diewert, W.E. (2018), "Scanner Data, Elementary Price Indexes and the Chain Drift Problem", Discussion Paper 18-06, Vancouver School of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1L4.

Diewert, W.E (2020), "The Chain Drift Problem and Multilateral Indexes", Chapter 7, *CPI Manual*, ILO-ECE, Geneva. *Forthcoming.*

Dynan, Karen, E. 2000. "Habit Formation in Consumer Preferences: Evidence from Panel Data." American Economic Review, 90 (3): 391-406.

Eltetö, Ö., and Köves, P. (1964), "On a Problem of Index Number Computation Relating to International Comparisons", (in Hungarian), Statisztikai Szemle 42, 507-518.

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

Feenstra, R.C. and M.D. Shapiro (2003), "High-Frequency Substitution and the Measurement of Price Indexes", pp. 123-146 in Scanner Data and Price Indexes, Robert C. Feenstra and Matthew D. Shapiro (eds.), Studies in Income and Wealth Volume 64, Chicago: The University of Chicago Press.

Feenstra, R.C. and Xu, M., and A. Antoniades, (2019), "What is the Price of Tea in China? Goods Prices and Availability in Chinese Cities." Available at SSRN: hiips://ssrn.com/abstract=3361233 or hiip://dx.doi.org/10.2139/ssrn.3361233

Geary, R.G. (1958), "A Note on Comparisons of Exchange Rates and Purchasing Power between Countries", Journal of the Royal Statistical Society Series A 121, 97-99.

Gini, C. (1931), “On the Circular Test of Index Numbers”, Metron 9:9, 3-24.

Hendel, I., & Nevo, A. (2006). Measuring the Implications of Sales and Consumer Inventory Behavior. Econometrica, 74(6), 1637-1673. Retrieved May 9, 2020,

ILO/IMF/OECD/UNECE/Eurostat/The World Bank (2020), Consumer Price Index Manual: Theory and Practice, Geneva: International Labour Office.

Ivancic, L., W.E. Diewert, and K.J. Fox (2009), “Scanner Data, Time Aggregation and the Construction of Price Indexes”, Discussion Paper 09-09, Department of Economics, University of British Columbia, Vancouver, Canada.

Ivancic, L., W. E. Diewert, and K. J. Fox (2011), “Scanner data, time aggregation and the construction of price indexes,” Journal of Econometrics, Volume 161, Issue 1, Pages 24-35.

Jevons, W.S. (1963) A Serious Fall in the Value of Gold, Edward Stanford.

Khamis, S.H. (1972), “A New System of Index Numbers for National and International Purposes”, Journal of the Royal Statistical Society Series A 135, 96-121.

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

Lamboray, C. (2017), ‘The Geary Khamis Index and the Lehr Index: How Much Do They Differ?’, paper prepared for the 15th Meeting of the Ottawa Group, 10-12 May 2017, Eltville.

Martin, R., S. (2019a), Changing tastes versus specification error in cost-of-living measurement, Working Paper, Bureau of Labor Statistics.

Martin, R. S. (2019b). “Revisiting taste change in cost-of-living measurement". Working Paper, Bureau of Labor Statistics

Neary, J. P., (2004), “Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income,” American Economic Review, 94, 1411–1428.

Osborn, D.R. (1988), Seasonality and habit persistence in a life cycle model of consumption. J. Appl. Econ., 3: 255-266. doi:10.1002/jae.3950030402

Redding, S. J. and D. E. Weinstein (2018). “Measuring Aggregate Price Indexes with Demand Shocks: Theory and Evidence for CES Preferences". NBER Working Paper

Redding, S. J. and D. E. Weinstein (2020), “Measuring Aggregate Price Indices with Taste Shocks: Theory and Evidence for CES Preferences", Quarterly Joural of Economics, 503–560. doi:10.1093/qje/qjz031

Reinsdorf, M. B., and A. H. Dorfman (1999), “The Sato-Vartia Indeex and the Monotonicity Axiom”, Journal of Econometrics, 90, 45-61.

Samuelson, P.A and S. Swamy (1974), “Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis”, The American Economic Review, Vol. 64, No. 4 (Sep., 1974), pp. 566-593

Sato, K., (1976) “The Ideal Log-Change Index Number,” Review of Economics and Statistics, 58, 223–228.

Szulc, B.J. (1964), “Indices for Multiregional Comparisons”, (in Polish), Przeglad Statystyczny 3, 239-254.

Szulc, B.J. (1983), “Linking Price Index Numbers,” pp. 537-566 in Price Level Measurement, W.E. Diewert and C. Montmarquette (eds.), Ottawa: Statistics Canada.

Szulc, B.J. (1987), “Price Indices below the Basic Aggregation Level”, Bulletin of Labour Statistics 2, 9-16.

Tornqvist, L. (1931, 36 37?)

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

von Auer, L. (2019), “The Nature of Chain Drift: Implications for Scanner Data Price Indices”, paper presented at the 16th Meeting of the Ottawa Group, Rio de Janeiro, Brazil, May 8.

Young, A.A. (1921) "The Measurement of Changes in of the General Price Level, " Quarterly Journal of Economics, Volume 35, Issue 4, August 1921, Pages 557–573,

APPENDIX

A1: Notation and Preliminaries

Let the vector of prices, quantities, and taste parameters for N commodities at time t be denoted by:

$$p_t = \left(p_{1t}, p_{2t}, ..., p_{Nt}\right),$$
$$q_t = \left(q_{1t}, q_{2t}, ..., q_{Nt}\right),$$
$$\varphi_t = \left(\varphi_{1t}, \varphi_{2t}, ..., \varphi_{Nt}\right).$$

The utitlity function at time t has the CES form as follows,

$$U_t\left(q_t;\varphi_t,\sigma\right) = \left(\sum_{i=1}^{N}\left(\varphi_{it}q_{it}\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \qquad \text{(A.1)}$$

The number of commodities is constant over time and fixed at $N>1$. $\sigma>1$ is the elasticity of substitution.

The unit cost function corresponding to utility function in (A.1) is

$$C\left(p_t,\varphi_t;\sigma\right) = \left(\sum_{i=1}^{N}\left(\frac{p_{it}}{\varphi_{it}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}.$$

The exogenous income, Y_t, is equal to the total expenditure at time t, that is, the following budget constraint holds for all the time periods,

$$\sum_{i=1}^{N} p_{it}q_{it} = Y_t.$$

Note that the following equation holds,

$$C\left(p_t,\varphi_t;\sigma\right)\times U_t\left(q_t;\varphi_t,\sigma\right) = \sum_{i=1}^{N} p_{it}q_{it}$$
$$= Y_t.$$

The weightes used in the Sato-Vartia index, ω_{ist}^{*}, is defined as

$$\omega_{ist}^{*} = \frac{w_{is}-w_{is}}{\ln\left(w_{it}\right)-\ln\left(w_{is}\right)} \Big/ \sum_{i=1}^{N}\left(\frac{w_{it}-w_{is}}{\ln\left(w_{it}\right)-\ln\left(w_{is}\right)}\right) \quad where \;\; w_{i\tau} = p_{i\tau}q_{i\tau}\Big/\sum\nolimits_{i=1}^{N} p_{i\tau}q_{i\tau} \quad i=1,2...,N;\ \tau=s,t$$

Note that the following equations hold for all s and t,

$$\sum_{i=1}^{N}\omega_{ist}^{*} = 1,$$

$$\sum_{i=1}^{N}\omega_{ist}^{*}\left(\ln w_{it}-\ln w_{is}\right)=0.$$

Appendix A1

Result 1 - Monotonicity of CUPI

We start with the result , for any $i=1,2,...,N$, we have

$$\ln CUPI_{st}=\ln P_t-\ln P_s$$

where

$$\ln P_t=-\ln\varphi_{it}+\ln p_{it}-\frac{1}{1-\sigma}\ln w_{it}$$
$$\ln P_s=-\ln\varphi_{is}+\ln p_{is}-\frac{1}{1-\sigma}\ln w_{is}.$$

Since P_s does not depend on p_{it}, to show the monotonicity, it is sufficient to show that P_t is increasing with p_{it}.

Because both w_{it} and φ_{it} are differentiable functions with respect to p_{it}, to establish monotonicity, it is sufficient to show the following inequality.

$$\frac{d\ln P_t}{d\ln p_{it}}=-\frac{d\ln\varphi_{it}}{d\ln p_{it}}+1-\frac{1}{1-\sigma}\frac{d\ln w_{it}}{d\ln p_{it}} \\ >0. \tag{A.2}$$

First, consider the third term of the R.H.S. of A. The derivative of $\ln w_{it}$ with respect to $\ln p_{it}$ is as follows,

$$\frac{d\ln w_{it}}{d\ln p_{it}}=\frac{d\ln\left(p_{it}q_{it}\right)}{d\ln p_{it}}-\frac{d\ln\left(\sum_{k=1}^{N}p_{kt}q_{kt}\right)}{d\ln p_{it}}$$
$$=1-\frac{q_{it}}{\sum_{k=1}^{N}p_{kt}q_{kt}}\times\frac{dp_{it}}{d\ln p_{it}}$$
$$=1-w_{it}.$$

Note that when $k\neq i$, we have

$$\frac{d\ln w_{kt}}{d\ln p_{it}} = -w_{kt}.$$

Next, consider the first term of the R.H.S. of (A.2). From the definition of the taste parameter, we have,

$$\varphi_{it} = \left(\frac{p_{it}}{p_{1t}}\right)\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\left[\prod_{k=2}^{N}\left\{\left(\frac{p_{kt}}{p_{1t}}\right)\left(\frac{w_{kt}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\right\}^{(-1/N)}\right]\varphi.$$

Taking natural logarithms leads us to

$$\ln\varphi_{it} = \frac{1}{\sigma-1}\ln\left(\frac{w_{it}}{w_{1t}}\right)+\ln\left(\frac{p_{it}}{p_{1t}}\right)-\left(\frac{1}{N}\right)\sum_{k=2}^{N}\left(\frac{1}{\sigma-1}\ln\left(\frac{w_{kt}}{w_{1t}}\right)+\ln\left(\frac{p_{kt}}{p_{1t}}\right)\right)+\ln\varphi. \quad \text{(A.3)}$$

The derivative of the first term of (A.3) with respect to $\ln p_{it}$ is

$$\frac{1}{\sigma-1}\frac{d\ln\left(\frac{w_{it}}{w_{1t}}\right)}{d\ln p_{it}} = \frac{1}{\sigma-1}\left(1-w_{it}+w_{1t}\right) \text{ if } \ i\neq 1$$
$$= 0 \text{ if } \ i=1.$$

The derivative of the second term is as follows,

$$\frac{d\ln\left(\frac{p_{it}}{p_{1t}}\right)}{d\ln p_{it}} = 1 \text{ if } \ i\neq 1$$
$$= 0 \text{ if } \ i=1.$$

The derivative of the third term when $i\neq 1$ can be written as follows,

$$\frac{d\left[\left(\frac{1}{N}\right)\sum_{k=2}^{N}\left(\frac{1}{\sigma-1}\ln\left(\frac{w_{kt}}{w_{1t}}\right)+\ln\left(\frac{p_{kt}}{p_{1t}}\right)\right)\right]}{d\ln p_{it}}$$
$$= \frac{1}{N}\sum_{k=2}^{N}\left(\frac{-w_{kt}+w_{1t}}{\sigma-1}\right)+\frac{1}{N}\left(\frac{1}{\sigma-1}\right)+\left(\frac{1}{N}\right)$$
$$= \frac{1}{N(\sigma-1)}(N-1)(w_{1t})-\frac{1}{N(\sigma-1)}\sum_{k=2}^{N}(w_{kt})+\left(\frac{1}{N}\right)+\frac{1}{N}\left(\frac{1}{\sigma-1}\right)$$
$$= \frac{1}{N(\sigma-1)}(N-1)(w_{1t})-\frac{1}{N(\sigma-1)}(1-w_{1t})+\left(\frac{1}{N}\right)+\frac{1}{N}\left(\frac{1}{\sigma-1}\right)$$
$$= \frac{1}{N(\sigma-1)}\left(\sigma+Nw_{1t}-1\right).$$

Therefore, if $i \neq 1$

$$\frac{d \ln \varphi_{it}}{d \ln p_{it}}$$
$$= \frac{1}{\sigma-1}\left(1-w_{it}+w_{1t}\right)+1+\frac{1}{N\left(\sigma-1\right)}\left(\sigma+Nw_{1t}-1\right)$$
$$= \frac{1}{N\left(\sigma-1\right)}\left(\sigma+N\sigma+2Nw_{1t}-Nw_{it}-1\right)$$

Now, combining the derivatives of all the three terms, we can show that when $i \neq 1$

$$\frac{d \ln P_t}{d \ln p_{it}} = -\frac{d \ln \varphi_{it}}{d \ln p_{it}}+1-\frac{1}{1-\sigma}\frac{d \ln w_{it}}{d \ln p_{it}}$$
$$= \frac{-1}{N\left(\sigma-1\right)}\left(\sigma+N\sigma+2Nw_{1t}-Nw_{it}-1\right)+1-\frac{1}{1-\sigma}\left(1-w_{it}\right)$$
$$= \frac{1}{N\left(\sigma-1\right)}\left(\sigma+2Nw_{1t}-1\right)$$

Because $\sigma > 1,$ we obtain,

$$\frac{d \ln P_t}{d \ln p_{it}} > 0.$$

When $i = 1,$ the effects through the taste term can be obtained as follows:

$$\varphi_{1t} = \prod_{k=2}^{N}\left\{\left(\frac{p_{kt}}{p_{1t}}\right)\left(\frac{w_{kt}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\right\}^{(-1/N)} \qquad \varphi$$
$$\ln \varphi_{1t} = \frac{-1}{N}\left(\frac{1}{\sigma-1}\sum_{k=2}^{N}\ln\left(\frac{w_{kt}}{w_{1t}}\right)+\sum_{k=2}^{N}\ln\left(\frac{p_{kt}}{p_{1t}}\right)\right)$$
$$\frac{d \ln \varphi_{1t}}{d \ln p_{1t}} = \frac{-1}{N}\left(\frac{1}{\sigma-1}\sum_{k=2}^{N}\left(\left(w_{1t}-1\right)-w_{kt}\right)-1)\right)$$
$$= \frac{-1}{N}\left(\frac{1}{\sigma-1}\sum_{k=2}^{N}\left(w_{1t}-1\right)-\frac{1}{\sigma-1}\sum_{k=2}^{N}w_{kt}-1)\right)$$

$$= -\frac{1}{N\left(1-\sigma\right)}\left(N+\sigma-Nw_{1t}-1\right).$$

Note that we have

$$\frac{d\ln w_{1t}}{d\ln p_{1t}} = 1 - w_{1t}.$$

Therefore,

$$\frac{d\ln P_t}{d\ln p_{1t}} = -\frac{d\ln \varphi_{1t}}{d\ln p_{1t}} + 1 - \frac{1}{1-\sigma}\frac{d\ln w_{1t}}{d\ln p_{1t}}$$

$$= \frac{\sigma}{1-\sigma}\frac{1}{N\sigma}\left(N+\sigma-Nw_{1t}-1\right)+1-\frac{1}{1-\sigma}\left(1-w_{1t}\right)$$

$$= \frac{1}{N}\left(N-1\right) > 0.$$

This concludes the proof.

Appendix A2

Proof of Result 7 - Necessary and Sufficient condition for Commensurability

We prove this result using indiuction.

First, we prove when $N = 2$. Proof of this case is included in the main text but reproduced here for providing continuity of proof my mathematical induction.

From the definition of the taste parameter , we have.

$$\varphi_{2t} = \left(\frac{p_{2t}}{p_{1t}}\right)\left(\frac{w_{2t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\varphi_{1t}.$$

Suppose a change in measurement units occurs, so that we get a new price, p^*_{it}, and quantity, q^*_{ti}, as follows,

$$p^*_{it} = p_{it}\lambda_i,\ \ q^*_{it} = q^*_{it}/\lambda_i \quad \text{where } \lambda_i > 0 \text{ for } i = 1 \text{ or } 2.$$

Also denote φ^*_{1t} as the taste parameter for commodity i after the change in the measurement units. Then as is shown before, the necessary and sufficient condition for the commensurability is as follows,

$$\ln\varphi_{1s} - \ln\varphi^*_{1s} = \ln\varphi_{1t} - \ln\varphi^*_{1t}.$$

Denote the difference of the taste parameter as $\ln K$, that is,

$$\ln\varphi_{1s} - \ln\varphi^*_{1s} = \ln\varphi_{1t} - \ln\varphi^*_{1t} = \ln K \tag{A.4}$$

Also denote

$$x_t = \left(\frac{p_{2t}}{p_{1t}}\right)\left(\frac{w_{2t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}$$
$$x_t^* = \left(\frac{p_{2t}^*}{p_{1t}}\right)\left(\frac{w_{2t}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}$$
$$\lambda = \left(\frac{\lambda_2^*}{\lambda_1}\right).$$

By assumption, the normalization condition before change in units of measurement is

$$\varphi_{1t} = g(x_t).$$

Then, after the change in the measurement units, we have

$$\varphi_{1t} = g(x_t^*) = g(\lambda x_t).$$

Note that (A.4) can be written as

$$g(\lambda x_t) - g(x_t) = g(\lambda x_s) - g(x_s) = \ln K.$$

The above equation must hold for any values for x_t and x_s. Therefore, while K depends on λ, K does not depend on x_t nor x_s.

If the index passes the commensurability test, given λ, there is a constant $K(\lambda) > 0$ for all x_t that satisfy the following equations.

$$g(x_t) = \frac{1}{K(\lambda)} g(\lambda x_t).$$

Suppose $x_t = 1$, then, $K(\lambda)$ can be written as the ration between $g(\lambda)$ and $g(1)$ as follows,

$$g(1) = \frac{1}{K(\lambda)} g(\lambda)$$
$$g(\lambda) = K(\lambda) g(1)$$
$$K(\lambda) = \frac{g(\lambda)}{g(1)}.$$

Therefore, we can obtain the following equation for all $x_t > 0$ and $\lambda > 0$.

$$g(x_t) g(\lambda) = g(1) g(\lambda x_t). \tag{A.5}$$

For any $y > 0$, define a function $f : \mathbf{R}_{++} \to \mathbf{R}_{++}$,

$$f(y) = g(y)\sqrt{g(1)},$$

then, (A.5) can be written as

$$f(x_t)\, f(\lambda) = f(\lambda x_t). \tag{A.6}$$

Equation (A.6) is one of the classical Cauchy's functional equations whose general non-zero solution for $x > 0$ and $\lambda > 0$ is given by[22]

$$f(y) = y^c, \quad \text{with } c \in \mathbf{R}.$$

Therefore, we have

$$\begin{aligned} g(x) &= \frac{x^c}{\sqrt{g(1)}} \\ &= ax^c \quad c \in \mathbf{R},\, a \in \mathbf{R}_{++}. \end{aligned}$$

Therefore, $g(x)$ must be of the following form,

$$\varphi_{1t} = ax_{t2}^c.$$

Next, we consider a general case with $N \geq 3.$ Suppose we have the following normalization condition,

Using the definition of the taste parameters, the normalization condition can be written as

$$\varphi_1 = g(x_{2t}, x_{3t}, ..., x_{Nt}).$$

where

$$x_{it} = \left(\frac{p_{it}}{p_{1t}}\right)\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}.$$

After the change in the measurement units, we have

$$\varphi_{1t}^* = g(x_{2t}^*, x_{3t}^*, ..., x_{Nt}^*) = g(\pi_2 x_{2t}, \pi_3 x_{3t}, ..., \pi_N x_{Nt})$$

where

$$\pi_i = \left(\frac{\lambda_i}{\lambda_1}\right).$$

As is shown before, the necessary and sufficient condition for the commensurability is

$$\ln \varphi_{1s}^* - \ln \varphi_{1s} = \ln \varphi_{1t}^* - \ln \varphi_{1t}.$$

The above condition can be written as,

$$\ln g(x_{2t}^*, x_{3t}^*, ..., x_{Nn}^*) - \ln g(x_{2t}, x_{3t}, ..., x_{Nt}) = \ln g(x_{2s}^*, x_{3s}^*, ..., x_{Ns}^*) - \ln g(x_{2s}, x_{3s}, ..., x_{Ns}).$$

[22] The derivation of this general solution for this functional equation can be found in various textbook on functional equation such as Aczel (1966) and Efthimious (2010)

Define a vector, π as

$$\pi = \left(\pi_2, \pi_3, ..., \pi_N\right).$$

Then, we can rewrite the above conditions as

$$\begin{aligned}
&= \ln g\left(\pi_2 x_{2s}, \pi_3 x_{3s}, ..., \pi_N x_{Ns}\right) - \ln g\left(x_{2s}, x_{3s}, ..., x_{Ns}\right) \\
&= \ln g\left(\pi_2 x_{2t}, \pi_3 x_{3t}, ..., \pi_N x_{Nt}\right) - \ln g\left(x_{2t}, x_{3t}, ..., x_{Nt}\right) \\
&\equiv K.
\end{aligned}$$

Since this condition must hold for any x_{ti} and x_{si}, K is independent from x_{ti} and x_{si}. That is, K is a function of π only. Therefore, it is possible to rewrite the necessary and sufficient condition as:

$$\begin{aligned}
\varphi_{1t} &= \frac{1}{K\left(\pi\right)} g\left(\pi_2 x_{2t}, \pi_3 x_{3t}, ..., \pi_N x_{Nt}\right) \\
&= g\left(x_{2t}, x_{3t}, ..., x_{Nt}\right)
\end{aligned}$$

Next, suppose

$$\left(x_{t2}, x_{t3}, ..., x_{tN}\right) = \left(1,1,....,1\right),$$

then, we have

$$\begin{aligned}
g\left(1,1,...,1\right) &= \frac{1}{K\left(\pi\right)} g\left(\pi_2, \pi_3, ..., \pi_N\right) \\
K\left(\pi\right) &= \frac{g\left(\pi_2, \pi_3, ..., \pi_N\right)}{g\left(1,1,...,1\right)}.
\end{aligned}$$

Therefore, for CUPI to be commensurable, for any $\pi > 0$ and $x_t > 0$, we must have

$$g\left(x_{2t}, x_{3t}, ..., x_{Nt}\right) g\left(\pi_2, \pi_3, ..., \pi_N\right) = g\left(1,1,...,1\right) g\left(\pi_2 x_{2s}, \pi_3 x_{3s}, ..., \pi_N x_{Ns}\right). \tag{A. 7}$$

Next, suppose

$$\begin{aligned}
x_{2t} &= x_{3t} = .. = x_{i-1t} = x_{i+1t} = ... = x_{Nt} = 1, \\
\pi_2 &= \pi_3 = .. = \pi_{i-1} = \pi_{i+1} = ... = \pi_N = 1.
\end{aligned}$$

then, (A.7) becomes

$$g\left(1,1,..,1, x_{it}, 1,..,1\right) g\left(1,1,..,1, \pi_i, 1,..,1\right) = g\left(1,1,...,1\right) g\left(1,1,.,1, x_{it}\pi_i, 1,..,1\right).$$

In the case with $N = 2$, we discussed that the general solution of this Cauchy's functional equation for $x_{ti} > 0$ and $\pi > 0$ is as follows,

$$g\left(1,1,..,1, x_{it}, 1,..,1\right) = x_{it}^{c_i} \times A.$$

Denote A as

$$A = \prod_{i=2}^{N} a_i,$$

then, for we have

$$g\left(1,1,..,1,x_{ti},1,..,1\right) = x_{it}^{c_i} \prod_{i=2}^{N} a_i$$

Now, suppose all x_{tk} for $k \geq i$ are unity, that is,

$$x_{kt} = x_{k+1t} = ... = x_{Nt} = 1.$$

Also, suppose that the following functional equation,

$$g\left(x_{2t}, x_{3t}, ., x_{k-1t}, 1, .., 1\right) g\left(\pi_2, \pi_3, ., \pi_{k-1}, 1, .., 1\right) = g\left(1,1,...,1\right) g\left(\pi_2 x_{2t}, \pi_3 x_{3t}, .., \pi_{k-1} x_{k-1t}, 1, ..., 1\right),$$

has the following general solution,

$$g\left(x_{2t}, x_{3t}, ., x_{k-1t}, 1..,1\right) = \left(\prod_{i=2}^{k-1} x_{it}^{c_i}\right)\left(\prod_{i=2}^{N} a_i\right).$$

Note that $g\left(1,1,..,1\right)$ is given by,

$$g\left(1,1,..,1\right) = \left(\prod_{i=2}^{N} a_i\right).$$

Applying the above result, set

$$g\left(x\right) = g\left(x_{2t}, x_{3t}, ., x_{k-1t}, 1, .., 1\right)$$
$$g\left(y\right) = g\left(1,1,..,1,x_{kt},1..,1\right),$$

then, we have

$$g\left(1,1,..,1,x_{tk},1..,1\right) \times g\left(x_{2t}, x_{3t}, ., x_{k-1t}, 1, .., 1\right)$$

$$= \left(\prod_{i=1}^{k} x_{it}^{c_k}\right)\left(\prod_{i=2}^{N} a_i^2\right)$$

$$= g\left(x_{2t}, x_{3t}, ., x_{k-1t}, x_{kt}, .., 1\right) \times \left(\prod_{i=2}^{N} a_i\right).$$

Therefore,

$$g\left(x_{2t}, x_{3t}, ., x_{k-1t}, x_{kt}, .., 1\right) = \left(\prod_{i=1}^{k} x_{it}^{c_k}\right)\left(\prod_{i=2}^{N} a_i\right).$$

By mathematical induction, the general solution of the functional equation of

$$g\left(x_{2t}, x_{3t}, ..., x_{Nt}\right) g\left(\pi_2, \pi_3, ..., \pi_N\right) = g\left(1,1,...,1\right) g\left(\pi_2 x_{2t}, \pi_3 x_{3t}, ..., \pi_N x_{Nt}\right),$$

is

$$g\left(x_{2t}, x_{3t}, ..., x_{Nt}\right) = \left(\prod_{i=1}^{N} x_{it}^{c_i}\right)\left(\prod_{i=2}^{N} a_i\right).$$

Therefore, to make CUPI be independent from the measurement unit of commodities, we must have the following normalization condition,

$$\varphi_1 = A\times\left(\prod_{i=2}^{N} x_{it}^{c_i}\right)\left(\prod_{i=2}^{N} a_i\right)$$

where

$$A = \left(\prod_{i=2}^{N} a_i\right) > 0 \qquad c_i \in \mathbf{R} \text{ for all } i \geq 2.$$

Note that we have

$$\varphi_{it} = \left(\frac{p_{it}}{p_{1t}}\right)\left(\frac{w_{it}}{w_{1t}}\right)^{\frac{1}{\sigma-1}}\varphi$$
$$= x_{it}\varphi_{1t}.$$

Therefore, the normalization condition can be written as

$$\varphi_{1t} = A\times\left(\prod_{i=2}^{N} x_{it}^{c_k}\right)$$
$$= A\times\left(\prod_{i=2}^{N}\left(\frac{\varphi_{it}}{\varphi_{1t}}\right)^{c_i}\right)$$

Denote

$$\sum_{i=x}^{N} c_i = \bar{c},$$

then, we get

$$\varphi_{1t} = A\times\left(\prod_{i=2}^{N}(\varphi_{it})^{c_i}\right)(\varphi_{1t})^{-\bar{c}}.$$

The above equation can be written as

$$A\times\left(\prod_{i=2}^{N}(\varphi_{it})^{c_i}\right)(\varphi_{1t})^{-\bar{c}-1} = 1.$$

Now consider the case where all commodities are treated similarly, then

$$c_i = \beta \text{ for all } i \geq 2$$

Then, the above equation becomes

$$A\times\left(\prod_{i=2}^{N}(\varphi_{it})^{\beta}\right)(\varphi_{1t})^{-(N-1)\beta-1} = 1.$$

Then, we find β that satisifies the following equation.

$$\beta = -\left(N-1\right)\beta - 1.$$

The solution is

$$\beta = \frac{-1}{N}.$$

Using β,

$$A \times \left(\prod_{i=2}^{N} (\varphi_{it})^{\beta}\right) (\varphi_{1t})^{\beta} = 1$$

Take

$$\left(\prod_{i=2}^{N} (\varphi_{it})^{\beta}\right)^{1/\beta} (\varphi_{1t})^{\beta} (A)^{1/\beta} = 1$$

we get

$$\left(\prod_{i=2}^{N} (\varphi_{it})\right) (\varphi_{1t}) (A)^{1/\beta} = 1$$

Therefore, we get

$$\prod_{i=1}^{N} (\varphi_{it}) = \kappa,$$

where

$$\kappa = (A)^{-1/\beta} > 0$$

End or Proof.

Appendix A3

Table A1: Some Descriptive Statistics of CUPI, S-V, and Discrepancies

	(CUPI-S-V)/S-V	DCUPI	DChained S-V	DS-V
N	143	142	142	142
Mean	0.003	-0.046	-0.139	-0.041
Median	0.003	-0.065	-0.066	0.028
Standard Deviation	0.005	0.911	0.772	0.817
Min	-0.015	-2.573	-2.295	-2.473
Max	0.014	2.259	2.034	1.850

Note: DCUPI, Dchained S-V, and DS-V are weekly change rates (percentages) from the previous weeks of CUPI, Chained Sato-Vartia, and Direct Sato-Vartia index, respectively.
Because CUPI is transitive, the selection of base does not affect the change
(CCV-S-V)/S-V is the discrepancies between the level of CUPI and direct Sato-Vartia whose bases are the first week of the sample period.