No. DP26-11

RCESR Discussion Paper Series

Commensurability in Hedonic Price Indices: Axiomatic Foundations and The Necessity of Log Transformations

April 2026

Naohito Abe, Hitotsubashi University

And

DS Prasada Rao

RCESR

The Research Center for Economic and Social Risks
Institute of Economic Research
Hitotsubashi University

2-1 Naka, Kunitachi, Tokyo, 186-8603 JAPAN
http://risk.ier.hit-u.ac.jp/

Commensurability in Hedonic Price Indices: Axiomatic Foundations and The Necessity of Log Transformations

Naohito Abe* and DS Prasada Rao†

April, 2026

Abstract

Accurate measurement of price change over time and across countries is central to macroeconomic analysis and welfare assessment. The increasing availability of scanner and web-scraped data means that price comparisons must be constructed from unbalanced data where the set of goods varies across time and space. A central requirement for meaningful price comparisons is commensurability: bilateral comparisons should remain invariant to changes in units of measurement. This paper develops an axiomatic foundation for regression-based price comparisons under incomplete commodity coverage and establishes an analytical framework for examining the commensurability requirement in time-product-dummy and time-dummy hedonic models. We show that the commensurability axiom imposes strong restrictions on the functional form of hedonic price models. In particular, the logarithmic transformation is necessary for commensurability. In CPD/TPD-type models, and more generally in pooled hedonic regressions with commodity fixed effects, the log specification is both necessary and sufficient, as rescaling shocks are absorbed by item effects. By contrast, in hedonic models without commodity fixed effects, the log transformation remains necessary but is generally insufficient in unbalanced panels, changes in unit of measurement tend to induce distortions. The analysis extends to imputation methods. Pooled first-stage imputation can preserve commensurability when a common item-specific absorption channel is present, whereas period-by-period imputation typically does not. This creates a trade-off: time-varying specifications enhance flexibility but weaken unit invariance. Closed-form expressions are derived to characterise the role of unbalancedness in time dummy

*Hitotsubashi University, nabe@ier.hit-u.ac.jp

†University of Queensland, d.rao@uq.edu.au

We thank comments from Chihiro Shimizu, Hideyuki Mizobuchi, Valentin Zelenyuk and seminar participants at Hitotsubashi University for their helpful comments. This research was supported by the Joint Usage/Research Center at the Institute of Economic Research, Hitotsubashi University (IERPK2315, IERPK2405), and by MEXT/JSPS KAKENHI (23H00799, 24H00012). The authors extend their thanks to Intage for facilitating the use of scanner data.

hedonic models. Empirical results using Japanese beverage scanner data show that time-dummy hedonic indices can exhibit substantial departures from commensurability even in direct temporal comparisons, while imputation-based Jevons indices display smaller bilateral effects that can nevertheless accumulate into chain drift. These findings suggest that hedonic and imputation-based methods are best suited to relatively homogeneous product groups where unit comparability can be controlled, and they highlight the need for explicit commensurability checks in applied price measurement.

1 Introduction

Accurate measurement of price change over time and across countries is central to macroeconomic analysis and policymaking. In temporal contexts, price indices underpin inflation-targeting frameworks, cost-of-living adjustments International Labour Office (2017), and the monitoring of inflation pressures by central banks using CPI and related measures, including trimmed indicators; see also Board of Governors of the Federal Reserve System (2025). In cross-country contexts, price comparisons are essential for assessing material well-being, poverty, and inequality, as discussed by Ravallion et al. (1991), Deaton (2010), and Asian Development Bank (2020). At the same time, the growing importance of housing in household budgets has highlighted the need for improved measurement of rents and house price dynamics.[1] Both CPI and purchasing power parity (PPP) measures are constructed through multi-stage aggregation procedures that combine detailed price information into higher-level indices, relying on the assumption that meaningful comparisons can be derived from observed price data even when the underlying commodity space is large and heterogeneous, as discussed in International Labour Office et al. (2020), Rao (2013), Asian Development Bank (2020), and Feenstra et al. (2015).

Given these wide-ranging applications, even relatively small differences in price measurement can have large implications for policy and for the assessment of economic outcomes; as emphasised by Deaton (2010), modest revisions in international price comparisons can lead to substantial changes in global poverty estimates. Recent growth in scanner and web-scraped price data has brought renewed attention to a long-standing difficulty in index number measurement: the set of goods and services varies substantially across time and across countries, as emphasized by Broda and Weinstein (2010) and Handbury (2021). This issue is particularly acute in markets characterised by rapid product turnover, entry and exit, and quality change, such as consumer electronics and other differentiated goods, as discussed by Braun and Lein (2021) and Diewert and Shimizu (2025). When commodity sets differ, standard index-number formulas such as Laspeyres, Paasche, and Fisher face a basic difficulty: they are defined on matched price vectors, yet matched vectors are precisely what unbalanced panels do not provide. When panels are unbalanced, some form of completion or "bridging" of missing observations is unavoidable, either explicitly through imputation or implicitly through modelling assumptions.

[1]See, for example, Hill (2013) and Hill and Syed (2016).

Two broad regression-based approaches dominate applied work: dummy-variable methods and hedonic (or imputation-based) methods. A convenient starting point for both is the time-dummy hedonic (TDH) specification, in which transformed prices are regressed on observed characteristics together with time or country dummies, and bilateral comparisons are inferred from the estimated dummy component. In this sense, TDH provides a general framework for quality-adjusted price comparison. The country-product-dummy (CPD) and time-product-dummy (TPD) models widely used in PPP and CPI practice can be viewed as special cases of this framework in which the characteristic component is saturated by product-specific effects, as discussed by Triplett (2004) and de Haan (2010).

The second approach relies on hedonic regression and imputation. In this case, a hedonic model is estimated on the observed data and then used to predict missing prices, thereby completing the panel and enabling index construction, as in Pakes (2003) and Erickson and Pakes (2011). More recent contributions have expanded the econometric and computational toolkit available for such first-stage models. For example, Bajari et al. (2025) use machine learning methods to extract product characteristics from high-dimensional data, while Calainho (2024) develops a flexible model-agnostic framework that allows nonlinear and nonparametric estimators. A practical advantage of imputation-based approaches is that they can be implemented period by period, allowing coefficients on product characteristics to vary over time and adapt to changing market conditions, as discussed by Silver and Heravi (2007) and Diewert and Shimizu (2026).

Despite the extensive use of these methods, relatively little attention has been paid to the choice of functional form for the dependent variable. In practice, researchers employ a variety of transformations, including levels, logarithms, and Box–Cox transformations, as discussed by Halvorsen and Pollakowski (1981), Pakes (2003), Erickson and Pakes (2011), and Box and Cox (1964). While the logarithmic specification is widely adopted for its interpretability and empirical performance, economic theory has provided limited guidance for this choice. As noted by Triplett (2004), specification decisions are often driven by convention rather than formal justification. This issue becomes particularly important when regression-based models are used to construct price indices for official statistics and policy analysis.

This paper addresses this gap by invoking the axiomatic property of commensurability: invariance of economically meaningful bilateral index numbers to changes in units of measurement of commodities. Commensurability formalises the requirement that a price comparison should not change merely because, for example, one commodity is expressed in "per kilogram" rather than "per gram." We demonstrate that this simple axiom has profound implications for model specification and the measurement of price change.

Our framework is built around a unified time-dummy hedonic specification, with CPD/TPD treated as a special case in which the non-time component spans the full item fixed-effect space. This unified formulation is important for two reasons. First, it allows commensurability to be defined at the level of bilateral comparison itself, rather than in terms of a model-specific coefficient normalisation. Second, it makes clear that the difference between CPD and hedonic models is not a difference in the definition of commensurability, but a difference in whether the rescaling shock can be absorbed by the non-time component of the specification.

A central message of the paper is that the commensurability axiom sharply separates dummy-variable methods from hedonic methods. In CPD/TPD-type models with commodity fixed effects, the logarithmic transformation is not only necessary but also sufficient for commensurability. In time-dummy hedonic models, by contrast, the logarithmic transformation is generally only a necessary condition: once the data are unbalanced, a change in the measurement unit of a single item typically propagates through re-estimation and changes the identified time or country comparisons themselves. Thus, even with log prices, hedonic methods are not generally unit-robust under incomplete commodity coverage. This implies an important practical limitation: hedonic methods should not be viewed as universally applicable index-number procedures for broad and heterogeneous product spaces, but rather as tools best applied to relatively homogeneous product groups where quality adjustment is the main objective and measurement units can be kept meaningfully comparable.

The main findings from our work are as follows. First, within the general class of time-dummy hedonic models, logarithmic transformation is a necessary condition for commensurability.[2] The proof uses the fact that CPD/TPD is nested within TDH: if a transformation fails commensurability even in the CPD special case, then it cannot be commensurable in the broader TDH class. Second, in the CPD/TPD special case, the logarithmic transformation is also sufficient, because unit changes can be fully absorbed by product-specific effects. Third, in unbalanced hedonic panels without such effects, logarithmic specifications are generally not sufficient: changes in measurement units propagate through re-estimation and alter the resulting bilateral comparisons. Fourth, in imputation-based approaches, the primary focus is on the vector of predicted prices. With a pooled first-stage regression containing a common item-specific absorption channel, affine-log specification ensures that imputed price relatives are independent of units of measurement. However, this result does not generally extend to period-by-period imputation: once the first-stage model is re-estimated separately in each period, the common absorption channel disappears, and the resulting predicted relatives are generally unit-sensitive in unbalanced designs. Thus, the very feature that makes period-specific imputation attractive in practice—its ability to accommodate time variation in attribute coefficients—also makes commensurability generally unattainable.

These results highlight a fundamental trade-off. Dummy-variable methods with product effects provide strong robustness to arbitrary unit changes but do not capture quality adjustment or innovation effects. Hedonic and imputation-based methods can incorporate such features and adapt to changing market conditions, but their resulting comparisons are generally sensitive to measurement units in the presence of incomplete data. Thus, the choice between these approaches is not only a question of econometric flexibility or

[2] Although their setting is not hedonic price measurement but the theory of the cost-of-living index, Konüs and Byushgens show that imposing commensurability on the index leads to a logarithmic form for the corresponding expenditure function. For an English translation of their paper and a useful discussion, see Diewert and Zelenyuk (2025).

predictive performance, but also a question of consistency with basic axiomatic principles.

Finally, we demonstrate that these issues are not only theoretical but also empirically relevant. Using weekly Japanese beverage scanner data, we find that the time-dummy hedonic index exhibits substantial departures from commensurability even at the level of direct time comparisons. By contrast, imputation-based Jevons indices display relatively small bilateral effects, but these small link-level discrepancies can accumulate through chaining into noticeable drift, especially under weekly imputation and in smaller stratified samples. Hence, the empirical relevance of commensurability depends not only on functional form and model structure, but also on how missing or sparse price data are handled and on whether comparisons are evaluated bilaterally or through chained aggregation.

Taken together, the paper makes four contributions. It provides an axiomatic justification for the logarithmic transformation in regression-based approaches to price index construction; offers a common framework for CPD/TPD and time-dummy hedonic models in which commensurability of the resulting price indices can be formalised; shows that the key issue associated with incomplete coverage is not only functional form but also whether the specification contains a time-invariant item-specific channel that can absorb the effects of changes in measurement units in both dummy-based regressions and imputation approaches; and, finally, demonstrates through illustrative examples and scanner-data evidence that departures from commensurability are not merely of theoretical consequence but also quantitatively relevant.

The remainder of the paper is organised as follows. Section 2 discusses the basic assumptions on data structure, including connected supports, the unified time-dummy specification, the recovery-based construction of bilateral comparisons, and the definition of commensurability. Section 3 investigates commensurability of price index numbers from time-dummy models, establishes logarithmic necessity in the general time-dummy hedonic class, and shows sufficiency in the CPD/TPD special case as well as in balanced hedonic panels. Section 4 focuses on unbalanced panels of price data and shows why the sufficiency property generally fails in these models, and derives closed-form propagation results in the no-attribute case under OLS and WLS. Section 5 turns to regression-based imputation, distinguishes pooled from period-by-period first-stage procedures, and characterises when predicted price relatives are invariant to single-good rescaling. Section 6 presents illustrations with simple numerical examples designed to demonstrate the theoretical results. Section 7 presents results from the analysis of Japanese beverage scanner data; the results in this section reinforce the theoretical findings discussed in Sections 3, 4, and 5. Section 8 concludes, and the Appendix collects the functional-equation results used in the proofs.

2 Preliminaries: Data Structure, Connected Supports, and Identification

There are two main regression-based approaches to price-index construction: (i) methods based on country/time product dummy models and (ii) methods based on imputation. This section introduces the basic framework used throughout the paper. In particular, the data structure, the role of connected supports, and the identifying assumptions on characteristics, common to both approaches, are described here.

2.1 Data Structure and Connected Supports

When working with commodity-level data, the observation pattern is often unbalanced. In scanner data, product turnover is frequent; in international comparisons, the set of available commodities may differ substantially across countries. In such environments, any coherent system of bilateral comparisons requires a basic overlap condition: observed prices must be connected through chains of common observations so that relative price information can propagate across the sample.

Without such overlapping links, the data split into disconnected "islands," and relative price levels across islands are not identified from the observed prices alone. In that case, any comparison across islands necessarily relies on additional structures not contained in the observed data.

Let $\Omega \subseteq \{1,\dots,N\} \times \{1,\dots,C\}$ denote the set of observed item–country (or item–time) cells. The panel may be unbalanced, and each cell (i,c) is assumed to contain at most one observed price P_{ic}.[3] We therefore begin by formalizing the minimal overlap condition under which relative price information can be propagated across observed cells.

Definition 2.1 (Connected data pattern)**.** Let $\Omega \subseteq \{1,\dots,N\} \times \{1,\dots,C\}$ denote the set of observed item–country cells for prices. Define the sets of observed items and countries (times) by

$$I^* := \{i \in \{1,\dots,N\} : (i,c) \in \Omega \text{ for some } c\}, \qquad J^* := \{c \in \{1,\dots,C\} : (i,c) \in \Omega \text{ for some } i\}.$$

We say that Ω is *connected* if for any two items $i, i' \in I^*$, there exist items $i = i_0, i_1, \dots, i_m = i'$ in I^* and countries $c_1, \dots, c_m \in J^*$ such that each consecutive pair of items is observed in a common country:

$$(i_{k-1}, c_k) \in \Omega \quad \text{and} \quad (i_k, c_k) \in \Omega \qquad \text{for } k = 1, 2, \dots, m. \tag{1}$$

Condition (1) states that any two items with observed prices can be linked through a chain of overlaps, where each consecutive pair of items is jointly observed in some

[3] When working with scanner data, items sometimes share the same product code but are sold in different stores. In such cases, one may either treat store–code combinations as distinct products or aggregate observations across stores within each code.

country (or period). Equivalently, in the associated bipartite item–country structure, the observed data form a connected component, so that any two countries (or periods) are also linked through chains of commonly observed items. In a balanced panel, the dataset is automatically connected. In unbalanced data, however, connectedness may fail when the set of observed items differs substantially across countries or periods. The following example illustrates such a disconnected pattern.

Example 1 (Disconnected pattern)**.** Suppose

$$\Omega = \{(1, A), (1, B), (2, C), (2, D)\}.$$

Then items 1 and 2 never co-appear in any country/time. The data split into two disjoint islands,

$$\{(1, A), (1, B)\} \qquad \text{and} \qquad \{(2, C), (2, D)\},$$

and bilateral comparisons are identified only within each island. Thus, while we can compare the price level between A and B, or between C and D, we cannot compare A with C or B with D without imposing additional structure beyond the observed prices.

2.2 A Regression Framework

We use a hedonic model that nests CPD/TPD as a special case. Let $\Omega \subseteq I \times T$ denote the observed item–time cells. For each $(i, t) \in \Omega$, consider

$$y_{it} = f(P_{it}) = \mu + h(z_i) + \delta_t + \varepsilon_{it}, \qquad (i, t) \in \Omega, \tag{2}$$

where $f : \mathbb{R}_{++} \to \mathbb{R}$ is strictly increasing with respect to P_{it}, the price of commodity i at time (country) t, $h(\cdot)$ is a possibly nonlinear characteristic component, δ_t is the time (country) effect, and ε_{it} is an additive error term. We assume throughout that the error term is invariant to changes in units of measurement: under a single-item rescaling experiment, the realizations ε_{it} remain unchanged, so that rescaling affects (2) only through the transformed dependent variable $y_{it} = f(P_{it})$.

Throughout, we assume that the characteristics are *time-invariant*, so that z_i does not depend on t. This assumption is common to the dummy-based and imputation-based settings studied below: it ensures that the non-time component captures persistent item heterogeneity rather than time variation in observed characteristics.[4]

We also assume that the characteristic component $h(z_i)$ is identifiable separately from the time dummies on each connected component of the observation pattern. That is,

[4]If item indicators were time-varying, the model would no longer preserve a stable non-time item component, and identification of time effects would generally fail. If observed characteristics were time-varying, the present notion of commensurability would also cease to be well-defined, since changes in transformed prices could then reflect changing characteristics rather than mere rescaling. In scanner data, this restriction is usually natural: products with different observed characteristics are typically assigned different product codes and treated as distinct items. The same difficulty arises if the parameters of the characteristic component vary over time, because the non-time component would then itself become time-dependent.

once the support is connected, the non-time component and the time effects are assumed not to be observationally confounded within that component. Our concern is not the identification problem for h itself, but the properties of the identified time comparisons under different units of measurements.

Model (2) includes CPD/TPD as a special case: if z_i is a saturated vector of item dummies and $h(z_i) = \alpha_i$, then (2) reduces to

$$y_{it} = \mu + \alpha_i + \delta_t + \varepsilon_{it}.$$

Thus, CPD/TPD is the special case of the unified model in which the non-time component spans the full item-fixed-effect space.

3 Dummy-Based Comparison and Commensurability

Section 2 introduced the common framework used throughout the paper: the observation pattern may be unbalanced, connected supports are required for identified comparisons, the characteristic variables are assumed to be time-invariant, and the non-time component is assumed to be identifiable separately from time effects on each connected component. We now introduce the dummy-based representation of bilateral comparison of general price levels and the corresponding notion of commensurability.

3.1 Normalization-invariant transformed levels

For later use, define the *identified transformed level* at time t by

$$\mu_t := \mu + \delta_t. \tag{3}$$

The value of δ_t itself depends on the normalization of the time dummies, whereas μ_t does not. Indeed, replacing (μ, δ_t) by $(\mu - v, \delta_t + v)$ leaves $\mu_t = \mu + \delta_t$ unchanged for every t. Hence μ_t, not δ_t itself, is the normalization-invariant transformed level identified by the model.[5]

3.2 Bilateral Comparison from Recovered Price Levels

Because μ_t is the normalization-invariant transformed level identified at time t, any bilateral comparison must ultimately be constructed from the collection $\{\mu_t\}$. At the same

[5]For example, one may normalize the time dummies by setting a particular reference period equal to zero, say $\delta_1 = 0$, or by imposing a zero-mean normalization such as $\sum_t \delta_t = 0$. These choices change the reported levels of the individual dummy coefficients δ_t, but they do not change the fitted quantity $\mu + \delta_t$. Thus the economically meaningful transformed level is the sum $\mu_t = \mu + \delta_t$, not the level of δ_t taken in isolation.

time, one must distinguish between two objects: the transformed level μ_t identified by the regression, and the underlying price level that the comparison is intended to represent.

This distinction is central to the interpretation of hedonic estimation. The regression does not directly decompose prices on the original scale; rather, it decomposes transformed prices into a non-time component and a time effect. In the dummy-based approach, the purpose is to remove the contribution of observable characteristics and residual variation so as to isolate the common time effect. The resulting object $\mu_t = \mu + \delta_t$ is therefore a transformed level, not yet a price level on the original scale.[6]

If μ_t is interpreted as the transformed image of an underlying price level P_t, so that

$$\mu_t = f(P_t),$$

then recovering P_t requires undoing the transformation used in estimation. Since f is strictly increasing, it is invertible on its image $I := f(\mathbb{R}_{++})$, and the recovery mapping is therefore naturally

$$\phi := f^{-1}.$$

Thus $\phi(\mu_t)$ is the price level recovered from the identified transformed level μ_t.

The key point is that the regression identifies only a scalar for each period: the recovered price level. Accordingly, any bilateral comparison in the present framework is ultimately a rule for comparing two positive scalars, namely the recovered levels $\phi(\mu_s)$ and $\phi(\mu_t)$. Because the framework is intended to apply to economies with an arbitrary positive number of commodities, this scalar comparison rule must remain coherent when the commodity space collapses to a single good. In a single-commodity economy, there is no remaining aggregation problem, and the most natural bilateral comparison between periods s and t is the ordinary price relative P_t/P_s.[7] This motivates the one-good consistency requirement in Proposition 3.1: whenever the economy reduces to a single commodity, the bilateral comparison function g should coincide with this price relative. The following proposition shows that this requirement uniquely determines the form of bilateral comparison.

Proposition 3.1 (One-good consistency implies the ratio form)**.** *Let $f : \mathbb{R}_{++} \to I \subseteq \mathbb{R}$ be continuous and strictly increasing, and let $\phi = f^{-1} : I \to \mathbb{R}_{++}$. Suppose that bilateral comparison is represented by a function $g : I \times I \to \mathbb{R}_{++}$. Assume that, whenever the economy contains only one commodity with prices $P_s, P_t > 0$ in periods s and t, and the identified transformed levels satisfy*

$$\mu_s = f(P_s), \qquad \mu_t = f(P_t),$$

[6] A parallel logic applies to imputation. There, the fitted hedonic structure is used to reconstruct missing item-level prices. Since the regression predicts transformed prices, recovering prices themselves again requires undoing the transformation used in estimation.

[7] In the one-good case, the ordinary price relative P_t/P_s also coincides with the cost-of-living index. Moreover, it satisfies essentially all of the standard axiomatic requirements usually imposed on bilateral price indices in such a setting, including identity, time reversal, transitivity, commensurability, factor reversal, and homogeneity of degree zero with respect to a common proportional change in P_s and P_t.

the bilateral comparison coincides with the ordinary price relative:

$$g(\mu_s, \mu_t) = \frac{P_t}{P_s}.$$

Then necessarily

$$g(\mu_s, \mu_t) = \frac{\phi(\mu_t)}{\phi(\mu_s)} \qquad \textit{for all } \mu_s, \mu_t \in I. \tag{4}$$

Proof. Because $\phi = f^{-1}$, we have

$$P_s = \phi(\mu_s), \qquad P_t = \phi(\mu_t).$$

Hence the one-good consistency requirement implies

$$g(\mu_s, \mu_t) = \frac{P_t}{P_s} = \frac{\phi(\mu_t)}{\phi(\mu_s)}.$$

Since f is continuous and strictly increasing, its image is an interval, and every pair $(\mu_s, \mu_t) \in I \times I$ can be written as $\mu_s = f(P_s)$ and $\mu_t = f(P_t)$ for some $P_s, P_t > 0$. Thus the conclusion holds for all $\mu_s, \mu_t \in I$. □

Proposition 3.1 shows that one-good consistency uniquely selects the ratio form. Motivated by this result, we adopt the following definition.[8]

Definition 3.1 (Recovery-based comparison)**.** Let $f : \mathbb{R}_{++} \to I \subseteq \mathbb{R}$ be continuous and strictly increasing, and let $\phi := f^{-1} : I \to \mathbb{R}_{++}$ denote its inverse. We say that a bilateral comparison is *recovery-based* if it is constructed from the identified transformed levels by

$$g(\mu_s, \mu_t) = \frac{\phi(\mu_t)}{\phi(\mu_s)} \qquad \text{for all } \mu_s, \mu_t \in I. \tag{5}$$

Thus, the bilateral comparison function $g(\mu_s, \mu_t)$ can be interpreted as the price index comparing price levels in s and t.

3.3 Necessity and Sufficiency of Commensurability

We now introduce commensurability. Its intuition is that a change in the measurement unit of a single item is a matter of recording convention and should therefore leave the economically meaningful bilateral price comparison unchanged.

Given an item i_0 and a scalar $K > 0$, consider the single-item rescaling experiment

$$P_{i_0 t} \mapsto K\, P_{i_0 t} \quad \text{for all } t \in T_{i_0}, \qquad P_{it} \text{ unchanged for } i \neq i_0,$$

[8] It is worth noting that Equation (5) implies that g is transitive. This does not mean, however, that the present framework coincides with the classical aggregation problem studied by Funke et al. (1979). Here, bilateral comparison is constructed from the ratio of two recovered one-dimensional transformed levels, μ_s and μ_t, rather than from an axiomatic aggregation of many commodity prices. Moreover, we do not impose proportionality as in Funke et al. (1979).

where

$$T_i := \{t : (i,t) \in \Omega\}.$$

Re-estimation is performed on the same connected pattern Ω. The formal definition is given as follows:

Definition 3.2 (Commensurability)**.** Fix a base country(period) 1. The time-dummy model (2) is *commensurable* if, for every item i_0 satisfying $|T_{i_0}| \geq 2$ within the connected component of period 1,[9] and for every $K > 0$, the single-item rescaling experiment leaves all identified bilateral comparisons unchanged. That is, if μ_t^{new} denotes the transformed level after re-estimation, then commensurability requires

$$g(\mu_s^{\text{new}}, \mu_t^{\text{new}}) = g(\mu_s, \mu_t) \qquad \text{for all } s, t \text{ in the connected component of } 1. \tag{6}$$

If Ω is disconnected, the definition is understood componentwise.

In Section 2, we introduced the common framework used throughout the paper: comparisons are built from the normalization-invariant transformed levels $\mu_t = \mu + \delta_t$, represented by a bilateral comparison (price index) function $g(\mu_s, \mu_t)$, and commensurability is defined as invariance of these bilateral comparisons under single-item rescaling.

We now apply the preceding framework to the time-dummy hedonic model. Our first result establishes the necessity of logarithmic function for the general time-dummy hedonic class by exploiting the fact that CPD/TPD is nested within it: if commensurability holds in the full TDH class, it must in particular hold in the CPD special case, so it is enough to study a balanced CPD benchmark. We then turn to sufficiency. In the CPD/TPD special case, the logarithmic transformation is sufficient even on unbalanced connected supports, whereas in the general hedonic model sufficiency can be established on balanced panels. The failure of this sufficiency property on unbalanced hedonic supports is deferred to Section 4.

Proposition 3.2 (Necessity of logarithmic transformation in the time-dummy hedonic model)**.** *Consider the exact time-dummy hedonic structure*

$$f(P_{it}) = \mu + h(z_i) + \delta_t, \qquad (i,t) \in \Omega, \tag{7}$$

with time-invariant characteristics z_i. Suppose that $f : \mathbb{R}_{++} \to I \subseteq \mathbb{R}$ is strictly increasing and that its inverse recovery map

$$\varphi := f^{-1} : I \to \mathbb{R}_{++}$$

is C^1 with $\varphi'(x) > 0$ on I. Suppose further that bilateral comparisons are represented by (5) *under Definition 3.1. If the model is commensurable in the sense of Definition 3.2, then*

$$f(p) = a \ln p + b \qquad (a > 0,\ b \in \mathbb{R}).$$

[9]This restriction ensures that the rescaled item is observed in at least two periods within the relevant connected component. If an item appears in only one period, rescaling it does not provide a meaningful test of whether the identified bilateral comparison across periods is invariant.

Proof. Because CPD/TPD is a special case of TDH, commensurability of the TDH model implies commensurability on any CPD design viewed as a special case of TDH. It is therefore sufficient to show that commensurability on such a CPD design forces f to be log-affine. Moreover, the rescaling experiment used below satisfies Definition 3.2, since the rescaled item is observed in both periods and therefore has $|T_{i_0}| \geq 2$. We therefore work on a balanced two-period CPD design and derive the logarithmic form directly from commensurability.

Accordingly, consider a balanced two-period support with $N = 2$ items, and restrict (7) to the CPD case

$$f(P_{it}) = \mu + \alpha_i + \delta_t.$$

Let

$$\mu_t := \mu + \delta_t.$$

Without loss of generality, normalize the item effects so that

$$\sum_{i=1}^{2} \alpha_i = 0.$$

Then, averaging the exact CPD relation over i gives

$$\mu_t = \frac{1}{2} \sum_{i=1}^{N} f(P_{it}).$$

Under Definition 3.1,

$$g(\mu_0, \mu_1) = \frac{\phi(\mu_1)}{\phi(\mu_0)}.$$

Now rescale one item k by $K > 0$ in both periods.

Because the data set are balanced, this induces shifts

$$\mu_t^{(K)} = \mu_t + \eta_t, \qquad \text{where} \quad \eta_t := \frac{1}{2}\big[f(KP_{kt}) - f(P_{kt})\big].$$

Commensurability implies

$$\frac{\phi(\mu_1 + \eta_1)}{\phi(\mu_0 + \eta_0)} = \frac{\phi(\mu_1)}{\phi(\mu_0)}.$$

Equivalently,

$$\frac{\phi(\mu_1 + \eta_1)}{\phi(\mu_1)} = \frac{\phi(\mu_0 + \eta_0)}{\phi(\mu_0)}.$$

Let the rescaled item be item 1, and write $\alpha_1 = \alpha$. Since $N = 2$ and $\alpha_1 + \alpha_2 = 0$, the exact CPD equations imply

$$f(P_{1t}) = \mu_t + \alpha, \qquad f(P_{2t}) = \mu_t - \alpha.$$

Hence

$$P_{1t} = \phi(\mu_t + \alpha).$$

Therefore

$$\eta_t = \frac{1}{2}\left[f\{K\phi(\mu_t + \alpha)\} - (\mu_t + \alpha)\right].$$

Substituting this expression into the preceding commensurability condition gives

$$\frac{\phi\left(\mu_1 + \frac{1}{2}\left[f\{K\phi(\mu_1 + \alpha)\} - (\mu_1 + \alpha)\right]\right)}{\phi(\mu_1)} = \frac{\phi\left(\mu_0 + \frac{1}{2}\left[f\{K\phi(\mu_0 + \alpha)\} - (\mu_0 + \alpha)\right]\right)}{\phi(\mu_0)}.$$

The two transformed levels μ_0 and μ_1 can be varied independently over nondegenerate intervals by varying the two-period CPD design while keeping α fixed. Hence, for every $K > 0$ and every admissible α, the function

$$H_{K,\alpha}(x) := \frac{\phi\left(x + \frac{1}{2}\left[f\{K\phi(x + \alpha)\} - (x + \alpha)\right]\right)}{\phi(x)}$$

must be constant in x.

By Lemma A.4, this functional equation can hold for all $K > 0$ and for all α in a nondegenerate interval only if

$$\phi(x) = Ae^{\lambda x}$$

for some $A > 0$ and $\lambda > 0$. Since $\phi = f^{-1}$, it follows that

$$f(p) = \frac{1}{\lambda}\ln p - \frac{\ln A}{\lambda}.$$

Thus

$$f(p) = a \ln p + b, \qquad a > 0, \quad b \in \mathbb{R}.$$

This proves necessity. □

A related observation applies to weighted least squares. In empirical applications, price comparisons are often estimated using weights based on sales, expenditures, or other measures of economic importance. If commensurability is required uniformly over all admissible positive weighting schemes, then the logarithmic necessity result is unchanged, because the equal-weight CPD benchmark is included as a special case. The next corollary records this implication.

Corollary 3.2.1 (Weighted least-squares class)**.** *Suppose commensurability is required for every balanced CPD benchmark estimated by WLS with strictly positive weights. Then*

$$f(p) = a \ln p + b, \qquad a > 0, \quad b \in \mathbb{R}.$$

Proof. Equal-weight OLS, $w_{it} \equiv 1$, is a member of this WLS class. Hence the assumption implies commensurability in the equal-weight balanced CPD benchmark. The conclusion therefore follows directly from Proposition 3.2. □

We now turn from necessity to sufficiency. The conclusions here are more delicate, because sufficiency depends not only on the logarithmic form of f but also on the extent to which the attribute variables can absorb the effect of a change in the measurement unit.

In the CPD/TPD special case, the non-time component spans the full item-fixed-effect space. As a result, the logarithmic transformation is sufficient for commensurability even on connected but unbalanced supports. In the general time-dummy hedonic model, by contrast, the non-time component is typically lower-dimensional. For that broader class, logarithmic sufficiency can still be established on balanced panels, where time comparisons are pinned down by within-item differences. Whether that sufficiency survives on unbalanced supports is a separate question, and its failure in general will be examined in Section 4.

Proposition 3.3 (Sufficiency in the CPD/TPD special case)**.** *Consider the exact CPD/TPD structure*

$$f(P_{it}) = \mu + \alpha_i + \delta_t, \qquad (i,t) \in \Omega,$$

on a connected, possibly unbalanced observation pattern Ω*, and suppose that bilateral comparisons are represented by* (5) *under Definition 3.1. If*

$$f(p) = a \ln p + b \qquad (a > 0,\ b \in \mathbb{R}),$$

then the model is commensurable in the sense of Definition 3.2.

Proof. Fix an item i_0 and a scalar $K > 0$, and rescale its observed prices by

$$P_{i_0 t} \mapsto K P_{i_0 t} \qquad \text{for all } t \in T_{i_0}.$$

Since $f(p) = a \ln p + b$, the transformed prices of item i_0 shift by the constant $a \ln K$:

$$f(K P_{i_0 t}) = f(P_{i_0 t}) + a \ln K \qquad \text{for all } t \in T_{i_0}.$$

Hence the rescaling can be absorbed exactly by redefining the item effect as

$$\alpha_{i_0}^{\text{new}} = \alpha_{i_0} + a \ln K, \qquad \alpha_i^{\text{new}} = \alpha_i \text{ for } i \neq i_0,$$

while keeping μ and all δ_t unchanged. Therefore the identified transformed levels

$$\mu_t = \mu + \delta_t$$

are unchanged for all t, and so the recovered bilateral comparisons

$$g(\mu_s, \mu_t) = \frac{\phi(\mu_t)}{\phi(\mu_s)}$$

are unchanged for all s, t in the connected component under consideration. Thus the model is commensurable. □

The CPD/TPD result relies on the fact that the non-time component spans the full item-fixed-effect space, so that a rescaling shock can be absorbed directly by the corresponding item effect. In the general time-dummy hedonic model, that exact absorption channel is not available in general. Nevertheless, on a balanced panel the logarithmic transformation is still sufficient, because time comparisons are then identified through within-item differences, and these differences are preserved under common rescaling of a given item across periods. The next proposition uses the same basic logarithmic-shift argument as Proposition 3.3. The difference is that, in the balanced hedonic setting, the shift need not be absorbed by a full item effect; instead, it cancels from within-item differences, which continue to identify time comparisons.

Proposition 3.4 (Sufficiency on balanced hedonic panels)**.** *Consider the exact time-dummy hedonic structure*

$$f(P_{it}) = \mu + h(z_i) + \delta_t, \qquad (i,t) \in I \times T,$$

on a balanced panel, where the characteristics z_i are time-invariant, and suppose that bilateral comparisons (price index functions) are represented by (5) *under Definition 3.1. If*

$$f(p) = a \ln p + b \qquad (a > 0,\ b \in \mathbb{R}),$$

then the model is commensurable in the sense of Definition 3.2.

Proof. Fix an item i_0 and a scalar $K > 0$. Under the logarithmic transformation, the rescaling changes the transformed prices of item i_0 by the same additive constant in every period:

$$f(KP_{i_0t}) = f(P_{i_0t}) + a \ln K \qquad (t \in T).$$

All other transformed prices are unchanged.

Because the panel is balanced and z_i is time-invariant, the time-effect contrasts are identified from within-item transformed price differences. The additional term $a \ln K$ is time-invariant for the rescaled item, and therefore cancels from every within-item time difference:

$$\{f(KP_{i_0t}) - f(KP_{i_0s})\} = f(P_{i_0t}) - f(P_{i_0s}).$$

Hence all identified time contrasts are unchanged:

$$\mu_t^{\text{new}} - \mu_s^{\text{new}} = \mu_t - \mu_s.$$

Since

$$\phi(x) = \exp\left(\frac{x-b}{a}\right),$$

the recovery-based comparison satisfies

$$g(\mu_s, \mu_t) = \frac{\phi(\mu_t)}{\phi(\mu_s)} = \exp\left(\frac{\mu_t - \mu_s}{a}\right).$$

Therefore

$$g(\mu_s^{\text{new}}, \mu_t^{\text{new}}) = \exp\left(\frac{\mu_t^{\text{new}} - \mu_s^{\text{new}}}{a}\right) = \exp\left(\frac{\mu_t - \mu_s}{a}\right) = g(\mu_s, \mu_t).$$

Thus all bilateral comparisons are unchanged, and the model is commensurable. □

Proposition 3.3 shows that in the CPD/TPD special case the logarithmic transformation is sufficient for commensurability even on unbalanced connected supports, because the rescaling shock is absorbed exactly by the item effect. Proposition 3.4 shows that in the general time-dummy hedonic model the same sufficiency conclusion continues to hold on balanced panels, where time comparisons are pinned down by within-item differences.

By contrast, once the support is unbalanced and the non-time component does not span the full item-fixed-effect space, this sufficiency property generally fails. The mechanism of that failure, and its implications for hedonic price comparison, are examined in detail in Section 4.

4 Unbalanced Hedonic Panels

Section 3 introduced the time/country dummy-based comparison framework: comparisons are built from the normalization-invariant transformed levels $\mu_t = \mu + \delta_t$, represented by a bilateral price index/comparison function $g(\mu_s, \mu_t)$, and commensurability is defined as invariance of these bilateral comparisons under single-item rescaling. Section 3 also showed that the logarithmic transformation is sufficient in the CPD/TPD special case even on unbalanced supports, and in the general time-dummy hedonic model on balanced panels. What remains is to understand why this sufficiency result for time-dummy hedonic models with balanced panels generally fails once the hedonic support is unbalanced.

We address that question in three steps. First, we give a counterexample showing that, even under the logarithmic transformation and time-invariant characteristics, unbalanced time-dummy hedonic models need not be commensurable. Second, we derive closed-form propagation formulas for the no-attribute case, first under OLS and then under WLS. Third, we explain why the same problem does not arise in CPD/TPD.

4.1 Setup

We continue to work with the time-dummy hedonic specification

$$f(P_{it}) = \mu + h(z_i) + \delta_t + \varepsilon_{it}, \qquad (i,t) \in \Omega, \tag{8}$$

where $f : \mathbb{R}_{++} \to \mathbb{R}$ is strictly increasing, z_i is time-invariant, and δ_t is the time effect. The error term ε_{it} is left unrestricted at this stage; the results below will impose additional conditions only when required by the particular estimation method under consideration.

The panel is now allowed to be unbalanced: the set of observed items may vary across periods. For each period t, let

$$N_t := \{i : (i,t) \in \Omega\}, \qquad T_i := \{t : (i,t) \in \Omega\}.$$

As before, the economically meaningful bilateral comparison is represented by

$$P_{st} = g(\mu_s, \mu_t), \qquad \mu_t := \mu + \delta_t,$$

and commensurability means invariance of these bilateral comparisons under a single-item rescaling experiment.

The key difference from the balanced case is that, in an unbalanced panel, the same item is not necessarily observed in every period. Hence the rescaling shock introduced by one item cannot in general be cancelled period by period through a common differencing argument. This is precisely where the CPD/TPD special case and the general hedonic case begin to diverge.

To exhibit a counterexample, we work here with the exact case $\varepsilon_{it} = 0$.

Proposition 4.1 (Failure of sufficiency in unbalanced hedonic panels)**.** *There exists an unbalanced time-dummy hedonic panel with time-invariant characteristics such that*

$$f(p) = a \ln p + b \qquad (a > 0)$$

and yet the model is not commensurable.

Proof. Consider two periods $t \in \{0, 1\}$ and three items A, B, C. Let the observed pattern be

$$\Omega = \{(A,0), (B,0), (C,0), (A,1)\},$$

so the panel is unbalanced: only item A appears in both periods.

We use the degenerate special case $h(z_i) \equiv 0$, which is nested in the general time-dummy hedonic specification with time-invariant characteristics. In particular, this specification contains no item-specific effect, so there is no channel through which a change in the measurement unit of item A can be absorbed; if $h(z_i) = \eta_i$, the model would instead reduce to the TPD/CPD case. Thus consider the logarithmic specification

$$\ln P_{it} = \mu + \delta_t.$$

Normalize $\delta_0 = 0$, and choose the original prices as in Table 1.

Under OLS, period 0 observations determine μ as their sample mean, while the period 1 observation determines $\mu + \delta_1$. Hence

$$\mu = \frac{0 + 1 + 2}{3} = 1, \qquad \mu + \delta_1 = 1,$$

Table 1: Original data in the unbalanced counterexample

Observation	$\boldsymbol{P_{it}}$	$\ln \boldsymbol{P_{it}}$
$(A,0)$	1	0
$(B,0)$	e	1
$(C,0)$	e^2	2
$(A,1)$	e	1

so

$$\mu = 1, \qquad \delta_1 = 0.$$

Now rescale item A by $K = e$ in both periods:

$$P_{A0} \mapsto e, \qquad P_{A1} \mapsto e^2.$$

The rescaled observed data are shown in Table 2.

Table 2: Observed data after rescaling item A by $K = e$

Observation	$\boldsymbol{P_{it}^{\text{new}}}$	$\ln \boldsymbol{P_{it}^{\text{new}}}$
$(A,0)$	e	1
$(B,0)$	e	1
$(C,0)$	e^2	2
$(A,1)$	e^2	2

Thus the rescaled transformed observations are

$$(1, 1, 2, 2).$$

Re-estimating the same model by OLS, period 0 observations now imply

$$\mu' = \frac{1+1+2}{3} = \frac{4}{3},$$

while the period 1 observation implies

$$\mu' + \delta_1' = 2.$$

Hence

$$\mu' = \frac{4}{3}, \qquad \delta_1' = \frac{2}{3}.$$

Therefore, before rescaling,

$$\mu_0 = \mu = 1, \qquad \mu_1 = \mu + \delta_1 = 1,$$

whereas after rescaling,

$$\mu_0' = \mu' = \frac{4}{3}, \qquad \mu_1' = \mu' + \delta_1' = 2.$$

For convenience, the identified transformed levels before and after rescaling are summarized in Table 3.

Table 3: Change in identified transformed levels under rescaling

	Before rescaling	**After rescaling**
μ_0	1	$\frac{4}{3}$
μ_1	1	2

Thus the transformed time contrast changes under the single-item rescaling:

$$\mu_1 - \mu_0 = 0, \qquad \mu_1' - \mu_0' = 2 - \frac{4}{3} = \frac{2}{3}.$$

Since $f(p) = \ln p$, the inverse transformation is

$$\phi(x) = e^x.$$

Hence the recovery-based bilateral comparison changes from

$$g(\mu_0, \mu_1) = \frac{\phi(\mu_1)}{\phi(\mu_0)} = \exp(\mu_1 - \mu_0) = 1$$

to

$$g(\mu_0', \mu_1') = \frac{\phi(\mu_1')}{\phi(\mu_0')} = \exp(\mu_1' - \mu_0') = e^{2/3}.$$

Therefore the bilateral comparison is not invariant under the single-item rescaling, and commensurability fails despite the logarithmic transformation. □

The counterexample above shows that once the dataset is unbalanced—as is often the case in practical hedonic estimation—price indices based on time-dummy hedonic models are generally not commensurable unless commodity fixed effects are included.

We now derive a general propagation formula that makes the source of this failure transparent.

4.2 Closed Form without Attributes: OLS

The propagation formula becomes especially transparent when there are no attributes.

Proposition 4.2 (Closed form without attributes under OLS)**.** *Consider the no-attribute* *time-dummy hedonic model*

$$y_{it} = a \ln P_{it} + b = \beta_0 + \delta_t + \varepsilon_{it}, \qquad (i,t) \in \Omega, \tag{9}$$

estimated by OLS on a possibly unbalanced panel, where $a > 0$, and impose the normalization $\delta_1 = 0$. Assume that under a single-item rescaling experiment, the error terms ε_{it} are unchanged, so that the rescaling affects the model only through the transformed dependent variable $y_{it} = a \ln P_{it} + b$.

Let

$$N_t := |\{i : (i,t) \in \Omega\}|, \qquad \bar{y}_t := \frac{1}{N_t} \sum_{i \in N_t} y_{it}.$$

Then

$$\hat{\beta}_0 = \bar{y}_1, \qquad \hat{\delta}_t = \bar{y}_t - \bar{y}_1 \qquad (t \neq 1).$$

Fix an item $i^\star$, and let

$$T_{i^\star} := \{t : i^\star \in N_t\}.$$

After rescaling item $i^\star$ by $K > 0$, the coefficient changes are

$$\hat{\beta}_0^{\text{new}} - \hat{\beta}_0 = a \ln K \cdot \frac{1}{N_1} \mathbf{1}_{\{i^\star \in N_1\}}, \tag{10}$$

$$\hat{\delta}_t^{\text{new}} - \hat{\delta}_t = a \ln K \left(\frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}} - \frac{1}{N_1} \mathbf{1}_{\{i^\star \in N_1\}} \right) \qquad (t \neq 1). \tag{11}$$

Equivalently, the identified transformed level

$$\hat{\mu}_t := \hat{\beta}_0 + \hat{\delta}_t$$

satisfies

$$\hat{\mu}_t^{\text{new}} - \hat{\mu}_t = a \ln K \cdot \frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}}. \tag{12}$$

Hence commensurability fails whenever

$$\frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}} \neq \frac{1}{N_s} \mathbf{1}_{\{i^\star \in N_s\}} \qquad \textit{for some pair } s, t. \tag{13}$$

In particular, relative to the base period 1, commensurability fails whenever

$$\frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}} \neq \frac{1}{N_1} \mathbf{1}_{\{i^\star \in N_1\}} \qquad \textit{for some } t \neq 1.$$

Proof. Under the normalization $\delta_1 = 0$, the model contains an intercept and time dummies for periods $t \neq 1$. Therefore OLS fits each period mean exactly. For the base period,

$$\hat{\beta}_0 = \bar{y}_1, \qquad \bar{y}_1 := \frac{1}{N_1} \sum_{i \in N_1} y_{i1}.$$

For each $t \neq 1$,

$$\hat{\beta}_0 + \hat{\delta}_t = \bar{y}_t, \qquad \bar{y}_t := \frac{1}{N_t} \sum_{i \in N_t} y_{it}.$$

Hence

$$\hat{\delta}_t = \bar{y}_t - \bar{y}_1, \qquad t \neq 1.$$

Now rescale item $i^\star$ by $K > 0$. Since

$$y_{it} = a \ln P_{it} + b,$$

this adds $a \ln K$ to the transformed dependent variable in exactly those observations for which $i = i^\star$. Because the error terms are unchanged by rescaling, no other part of the regression equation is affected.

For the base period, the period mean changes by

$$\bar{y}_1^{\text{new}} - \bar{y}_1 = a \ln K \cdot \frac{1}{N_1} \mathbf{1}_{\{i^\star \in N_1\}}.$$

Since $\hat{\beta}_0 = \bar{y}_1$, this gives

$$\hat{\beta}_0^{\text{new}} - \hat{\beta}_0 = a \ln K \cdot \frac{1}{N_1} \mathbf{1}_{\{i^\star \in N_1\}},$$

which proves (10).

Similarly, for each period t,

$$\bar{y}_t^{\text{new}} - \bar{y}_t = a \ln K \cdot \frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}}.$$

Therefore, for $t \neq 1$,

$$\begin{aligned} \hat{\delta}_t^{\text{new}} - \hat{\delta}_t &= (\bar{y}_t^{\text{new}} - \bar{y}_1^{\text{new}}) - (\bar{y}_t - \bar{y}_1) \\ &= a \ln K \left(\frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}} - \frac{1}{N_1} \mathbf{1}_{\{i^\star \in N_1\}} \right), \end{aligned}$$

which proves (11).

Finally, since

$$\hat{\mu}_t = \hat{\beta}_0 + \hat{\delta}_t = \bar{y}_t,$$

we have

$$\hat{\mu}_t^{\text{new}} - \hat{\mu}_t = \bar{y}_t^{\text{new}} - \bar{y}_t = a \ln K \cdot \frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}},$$

which proves (12). Hence, for any pair s, t,

$$(\hat{\mu}_t^{\text{new}} - \hat{\mu}_s^{\text{new}}) - (\hat{\mu}_t - \hat{\mu}_s) = a \ln K \left(\frac{1}{N_t} \mathbf{1}_{\{i^\star \in N_t\}} - \frac{1}{N_s} \mathbf{1}_{\{i^\star \in N_s\}} \right).$$

If this expression is nonzero for some pair s, t, then the identified time contrast changes under rescaling. Under the logarithmic transformation, the recovery-based comparison satisfies

$$g(\hat{\mu}_s, \hat{\mu}_t) = \exp\left(\frac{\hat{\mu}_t - \hat{\mu}_s}{a}\right).$$

Therefore the bilateral comparison changes, and commensurability fails. □

Note that condition (13) is knife-edge. It requires the rescaled item to contribute exactly the same share to every period mean relative to the base period. Any genuine unbalancedness in sample composition or sample size generically violates this condition.

The same logic extends to weighted least squares.

4.3 Closed Form without Attributes: WLS

The same propagation logic extends directly to weighted least squares.

Corollary 4.2.1 (Closed form without attributes under WLS)**.** *Consider the same no-attribute time-dummy hedonic model as in Proposition 4.2,*

$$y_{it} = a \ln P_{it} + b = \beta_0 + \delta_t + \varepsilon_{it}, \qquad (i,t) \in \Omega, \tag{14}$$

but now suppose estimation is by WLS with positive observation weights $\{w_{it} > 0 : (i,t) \in \Omega\}$, under the normalization $\delta_1 = 0$. Assume again that under a single-item rescaling experiment, the error terms ε_{it} are unchanged, so that the rescaling affects (14) only through the transformed dependent variable $y_{it} = a \ln P_{it} + b$.

Let

$$W_t := \sum_{i \in N_t} w_{it}, \qquad \bar{y}_t^w := \frac{1}{W_t} \sum_{i \in N_t} w_{it} y_{it}.$$

Then

$$\hat{\beta}_0 = \bar{y}_1^w, \qquad \hat{\delta}_t = \bar{y}_t^w - \bar{y}_1^w \qquad (t \neq 1).$$

Fix an item $i^\star$, and let

$$T_{i^\star} := \{t : i^\star \in N_t\}.$$

After rescaling item $i^\star$ by $K > 0$, the coefficient changes are

$$\hat{\beta}_0^{\text{new}} - \hat{\beta}_0 = a \ln K \cdot \frac{w_{i^\star 1}}{W_1} \mathbf{1}_{\{i^\star \in N_1\}}, \tag{15}$$

$$\hat{\delta}_t^{\text{new}} - \hat{\delta}_t = a \ln K \left(\frac{w_{i^\star t}}{W_t} \mathbf{1}_{\{i^\star \in N_t\}} - \frac{w_{i^\star 1}}{W_1} \mathbf{1}_{\{i^\star \in N_1\}} \right) \qquad (t \neq 1). \tag{16}$$

Equivalently, the identified transformed level

$$\hat{\mu}_t := \hat{\beta}_0 + \hat{\delta}_t$$

satisfies

$$\hat{\mu}_t^{\text{new}} - \hat{\mu}_t = a \ln K \cdot \frac{w_{i^\star t}}{W_t} \mathbf{1}_{\{i^\star \in N_t\}}. \tag{17}$$

Hence commensurability fails whenever

$$\frac{w_{i^\star t}}{W_t} \mathbf{1}_{\{i^\star \in N_t\}} \neq \frac{w_{i^\star s}}{W_s} \mathbf{1}_{\{i^\star \in N_s\}} \qquad \textit{for some pair } s, t. \tag{18}$$

In particular, relative to the base period 1*, commensurability fails whenever*

$$\frac{w_{i^\star t}}{W_t} \mathbf{1}_{\{i^\star \in N_t\}} \neq \frac{w_{i^\star 1}}{W_1} \mathbf{1}_{\{i^\star \in N_1\}} \qquad \textit{for some } t \neq 1.$$

Proof. The argument is the weighted analogue of Proposition 4.2. Under the normalization $\delta_1 = 0$, WLS fits each weighted period mean exactly, so

$$\hat{\beta}_0 = \bar{y}_1^w, \qquad \hat{\delta}_t = \bar{y}_t^w - \bar{y}_1^w \qquad (t \neq 1).$$

After rescaling item $i^\star$ by $K > 0$, the transformed dependent variable increases by $a \ln K$ exactly in those observations with $i = i^\star$. Therefore each weighted period mean changes by

$$\bar{y}_t^{w,\text{new}} - \bar{y}_t^w = a \ln K \cdot \frac{w_{i^\star t}}{W_t} \mathbf{1}_{\{i^\star \in N_t\}}.$$

Substituting this expression into $\hat{\beta}_0 = \bar{y}_1^w$ and $\hat{\delta}_t = \bar{y}_t^w - \bar{y}_1^w$ gives (15) and (16). Moreover, since

$$\hat{\mu}_t = \hat{\beta}_0 + \hat{\delta}_t = \bar{y}_t^w,$$

we obtain (17).

Finally, for any pair s, t,

$$(\hat{\mu}_t^{\text{new}} - \hat{\mu}_s^{\text{new}}) - (\hat{\mu}_t - \hat{\mu}_s) = a \ln K \left(\frac{w_{i^\star t}}{W_t} \mathbf{1}_{\{i^\star \in N_t\}} - \frac{w_{i^\star s}}{W_s} \mathbf{1}_{\{i^\star \in N_s\}} \right).$$

If this expression is nonzero, the identified time contrast changes. Since under $f(p) = a \ln p + b$,

$$g(\hat{\mu}_s, \hat{\mu}_t) = \exp\left(\frac{\hat{\mu}_t - \hat{\mu}_s}{a} \right),$$

the bilateral comparison changes as well. Hence commensurability fails. □

This failure is specific to hedonic specifications whose non-time component does not span the full item-fixed-effect space. It does not arise in CPD/TPD. There, the non-time component is a saturated system of item effects,

$$h(z_i) = \alpha_i,$$

so under the logarithmic transformation a single-item rescaling adds the same constant $a \ln K$ to all observed transformed prices of the rescaled item, and that shift is absorbed exactly by the corresponding item effect. Hence the identified transformed levels $\mu_t =$

$\mu + \delta_t$ remain unchanged, regardless of whether the observation pattern is balanced. Thus the unbalanced-panel failure of sufficiency is not a generic feature of dummy-based price models; it arises because the hedonic component spans only a lower-dimensional subspace, so the rescaling shock need not lie in that space.

5 Imputation and Commensurability

Having established the role of logarithmic transformations and item-specific absorption channels in dummy-based comparisons, we now examine whether the same logic carries over to regression-based imputation. We show that pooled imputation with commodity fixed effects preserves commensurability under the logarithmic specification, but that this property generally fails for period-by-period imputation, where the common absorption channel is lost.

The preceding sections studied time-dummy hedonic models, in which commensurability was formulated as invariance of the identified bilateral price comparison under changes in the unit of measurement. Although time-dummy-based index construction has long been used in practice, many official and empirical studies now also use regression-based imputation. This approach first estimates a model on the observed data, uses it to predict missing prices, and then constructs bilateral comparisons from the resulting observed and imputed price relatives.

This broad strategy includes linear and semiparametric hedonic imputation approaches such as Pakes (2003) and Erickson and Pakes (2011), as well as more recent machine-learning implementations for hedonic index construction Cafarella et al. (2023) and AI-based approaches that extract product characteristics from high-dimensional inputs Bajari et al. (2025). Despite differences in first-stage estimation, all such approaches ultimately rely on predicted prices to fill missing cells and to form bilateral comparisons. Under imputation, the primitive bilateral objects are therefore not dummy differences or transformed levels, but the *commodity-level price relatives* constructed from the combination of observed and predicted prices. Accordingly, the relevant commensurability requirement is invariance of those commodity-level relatives under unit of measurement change for single-good.

Imputation has an important practical attraction. Because it restores direct two-column comparisons at the commodity level when one side of the comparison is missing, it makes it possible to apply index-number formulae that are more tightly connected to economic index theory, including superlative index numbers in the sense of Diewert (1976) and other bilateral or multilateral constructions based on item-level price relatives. In this sense, imputation provides a bridge from regression-based quality adjustment to the broader theory of exact and superlative index numbers.

At the same time, this flexibility comes with a cost. Bilateral comparisons constructed from imputed item-level price relatives need not be transitive across periods, so under chaining they may be vulnerable to chain drift. By contrast, time-dummy approaches recover comparisons directly from a common regression system and therefore often provide a more tightly integrated comparison structure across periods. Thus, imputation and time-dummy approaches involve a genuine trade-off: the former is attractive because it permits the use of economically interpretable index-number formulae on a comparison set completed by prediction, whereas the latter is attractive because it builds the comparison directly into the regression-based time effects and may be less exposed to drift arising from chained bilateral imputations.

For this reason, commensurability under imputation deserves separate analysis. Even if a time-dummy specification yields a well-defined bilateral price comparison, it does not follow that the imputed commodity-level relative structure will be invariant under single-good rescaling; conversely, an imputation procedure may be attractive from the viewpoint of index-number theory while still inheriting unit sensitivity from the first-stage regression. The question in this section is therefore whether regression-based imputation preserves unit robustness at the level of the resulting commodity-level price relatives.

5.1 Commensurability under imputation

We first formalize the rescaling mechanism in a panel data setting, and then define commensurability directly in terms of the commodity-level relatives generated by the imputation rule.

Definition 5.1 (Single-good scaling operator)**.** Fix a commodity $i^\star$ and a factor $K > 0$. Given a possibly unbalanced panel $P = \{P_{ij}\}$ (with $j = c$ for countries or $j = t$ for time) and observed-cell set Ω, define the rescaled panel $S_{i^\star,K}(P)$ by multiplying the *observed* prices of commodity $i^\star$ by K:

$$(S_{i^\star,K}(P))_{ij} = \begin{cases} K\,P_{ij}, & (i,j) \in \Omega \text{ and } i = i^\star, \\ P_{ij}, & (i,j) \in \Omega \text{ and } i \neq i^\star, \\ \text{missing}, & (i,j) \notin \Omega. \end{cases}$$

Let $\widehat{P}_{ij}(P)$ denote the predicted *level* price produced by a fixed estimation-and-prediction procedure $\mathcal{E}$ whenever prediction is invoked. Fix a pair of columns (j_1, j_0), and define

$$I^\dagger := \big\{ i : \text{at least one of } (i, j_1), (i, j_0) \text{ is observed} \big\}.$$

For each $i \in I^\dagger$, define the imputation-based bilateral relative by

$$R_i(P) := \begin{cases} \dfrac{P_{ij_1}}{P_{ij_0}}, & \text{if both } (i,j_1) \text{ and } (i,j_0) \text{ are observed}, \\[2ex] \dfrac{\widehat{P}_{ij_1}(P)}{\widehat{P}_{ij_0}(P)}, & \text{if exactly one of } (i,j_1), (i,j_0) \text{ is observed}, \\[2ex] \text{undefined}, & \text{if neither } (i,j_1) \text{ nor } (i,j_0) \text{ is observed}. \end{cases}$$

Thus, if exactly one side of the two-column comparison is missing, both prices for that commodity are replaced by their predicted values; if both sides are observed, no imputation is performed.

Definition 5.2 (Commensurability under imputation)**.** Fix a base column j_0 and a comparison column j_1. We say that the estimation-and-prediction procedure $\mathcal{E}$ is *commensurable under imputation* for the pair (j_1, j_0) on a class of datasets $\mathcal{D}$ if for every $P \in \mathcal{D}$, every commodity $i^\star$, and every $K > 0$,

$$R_i\big(S_{i^\star,K}(P)\big) = R_i(P) \qquad \text{for all } i \in I^\dagger \text{ such that } R_i(P) \text{ is defined.}$$

Definition 5.2 is the natural analogue of commensurability for imputation. In the earlier sections, commensurability was formulated in terms of invariance of identified bilateral price comparisons built from transformed levels or dummy differences. Under imputation, by contrast, the primitive bilateral objects are the commodity-level relatives

$$\{R_i(P)\}_{i \in I^\dagger}.$$

Hence relative invariance, rather than coefficient invariance, is the relevant notion of unit robustness. Any bilateral index functional that depends only on this collection of relatives is automatically invariant whenever Definition 5.2 holds.

In the propositions below, $\mathcal{E}$ denotes the parametric estimation-and-prediction procedure generated by the stated model. On a connected observation pattern, fitted transformed values and the implied predicted prices are well defined up to the common location normalization.

5.2 Hedonic regressions with commodity fixed effects under imputation

We begin with the case in which the hedonic specification includes commodity-specific intercept shifts explicitly. This is the natural imputation analogue of the CPD/TPD benchmark studied earlier: the key issue is whether a single-good rescaling can be absorbed through an item-specific channel without altering the resulting commodity-level relatives.

Consider the *pooled* hedonic specification

$$y_{ij} = f(P_{ij}) = \mu + \alpha_i + \delta_j + \beta' z_{ij} + \varepsilon_{ij}, \tag{19}$$

estimated on a connected (possibly unbalanced) dataset Ω, where $f : \mathbb{R}_{++} \to \mathbb{R}$ is continuous and strictly increasing. We assume that the error term enters additively and is invariant under a single-good rescaling experiment: that is, rescaling the observed prices of one commodity by a factor $K > 0$ changes $f(P_{ij})$ only through the left-hand side of (19), while the realizations ε_{ij} themselves remain unchanged. For any cell (i, j), let

$$\widehat{P}_{ij}(P) := f^{-1}\big(\hat{\mu}(P) + \hat{\alpha}_i(P) + \hat{\delta}_j(P) + \hat{\beta}(P)' z_{ij}\big)$$

denote the fitted item-level prices used when imputation is implemented.

The next proposition shows that, once commodity fixed effects are included, commensurability under imputation holds exactly in the logarithmic case.

Proposition 5.1 (Commodity fixed effects restore logarithmic commensurability under imputation)**.** *Assume* (19) *is estimated on any connected (possibly unbalanced) dataset* Ω*. Then commensurability under imputation in the sense of Definition 5.2 holds for all single-good scalings if and only if*

$$f(p) = a \ln p + b \qquad \textit{with } a > 0.$$

Proof. **Necessity.** Consider the subclass $z_{ij} \equiv 0$, so that the model reduces to the TPD specification

$$f(P_{ij}) = \mu + \alpha_i + \delta_j + \varepsilon_{ij}.$$

It is enough to consider a two-period design. Take N commodities observed in both periods 0 and 1, and add one commodity i' observed only in period 0. This observation pattern is connected through period 0, and the N balanced commodities identify the period contrast.

Let $i^\star$ be one of the N balanced commodities. Rescale this commodity by $K > 0$ in both periods, and define

$$S_t(K) := f(KP_{i^\star t}) - f(P_{i^\star t}), \qquad t = 0, 1.$$

A direct two-way least-squares calculation on the balanced sub-block gives

$$\Delta(\hat{\delta}_1 - \hat{\delta}_0) = \frac{S_1(K) - S_0(K)}{N}.$$

The one-sided commodity i' is not rescaled. Since it is observed only in period 0, its fitted transformed value in period 0 remains equal to its observed transformed price. Hence, writing

$$x_t := \hat{\mu} + \hat{\alpha}_{i'} + \hat{\delta}_t,$$

we have

$$x_0^{new} = x_0, \qquad x_1^{new} = x_1 + \frac{S_1(K) - S_0(K)}{N}.$$

The imputation-based relative for commodity i' is

$$R_{i'}(P) = \frac{\phi(x_1)}{\phi(x_0)}.$$

Commensurability under imputation therefore requires

$$\frac{\phi\left(x_1 + \frac{S_1(K)-S_0(K)}{N}\right)}{\phi(x_0)} = \frac{\phi(x_1)}{\phi(x_0)}.$$

Since ϕ is strictly increasing, this implies

$$S_1(K) = S_0(K).$$

Thus, for arbitrary positive prices p_0, p_1,

$$f(Kp_1) - f(p_1) = f(Kp_0) - f(p_0).$$

Therefore, for each $K > 0$, the difference $f(Kp) - f(p)$ is independent of p. Hence there exists a function $A(K)$ such that

$$f(Kp) - f(p) = A(K) \qquad \text{for all } p > 0.$$

Setting $p = 1$ gives $A(K) = f(K) - f(1)$, so

$$f(Kp) = f(p) + f(K) - f(1).$$

Define

$$F(x) := f(e^x) - f(1).$$

Then

$$F(x + z) = F(x) + F(z).$$

By continuity, $F(x) = ax$ for some $a \in \mathbb{R}$. Since f is strictly increasing, $a > 0$. Consequently,

$$f(p) = a \ln p + f(1),$$

which is the required log-affine form. **Sufficiency.** Assume $f(p) = a \ln p + b$, with $a > 0$.

Fix any connected observation pattern Ω, which may be unbalanced, any commodity $i^\star$, and any $K > 0$. Under the single-good rescaling $P_{i^\star j} \mapsto KP_{i^\star j}$, the transformed dependent variable changes by

$$f(KP_{i^\star j}) - f(P_{i^\star j}) = a \ln K \qquad \text{for all } (i^\star, j) \in \Omega.$$

Let $(\hat{\mu}, \hat{\alpha}_i, \hat{\delta}_j, \hat{\beta})$ be a fitted solution for the original sample. Define

$$\hat{\alpha}_{i^\star}^{new} = \hat{\alpha}_{i^\star} + a \ln K, \qquad \hat{\alpha}_i^{new} = \hat{\alpha}_i \quad (i \neq i^\star),$$

and keep

$$\hat{\mu}^{new} = \hat{\mu}, \qquad \hat{\delta}_j^{new} = \hat{\delta}_j, \qquad \hat{\beta}^{new} = \hat{\beta}.$$

Then every fitted residual in the rescaled sample is identical to the corresponding fitted residual in the original sample. Hence, under the same estimation criterion, this shifted coefficient vector is a fitted solution for the rescaled sample. For commodities $i \neq i^\star$, neither observed prices nor fitted prices change. For commodity $i^\star$, observed prices are multiplied by K, and fitted transformed values are shifted by $a \ln K$. Since

$$\phi(x + a \ln K) = K\phi(x),$$

the fitted level prices for commodity $i^\star$ are also multiplied by K. Therefore, whether the relative is observed or imputed, both numerator and denominator for commodity $i^\star$ are multiplied by the same factor K, while all other commodity relatives are unchanged. Hence every defined commodity-level relative $R_i(P)$ is invariant. □

The TPD model is an immediate special case of this result. If the non-time component is a saturated system of item effects,

$$f(P_{ic}) = \mu + \alpha_i + \delta_c + \varepsilon_{ic},$$

then the specification contains an item-specific channel that absorbs exactly the additive shock generated by logarithmic single-good rescaling. Accordingly, TPD is commensurable under imputation for every comparison pair if and only if

$$f(p) = a \ln p + b \qquad (a > 0,\ b \in \mathbb{R}),$$

which is obtained from Proposition 5.1 by setting $z_{ij} \equiv 0$.

Thus, under imputation as in the earlier dummy-based analysis, what matters is not the label "TPD" versus "hedonic" as such, but whether the specification contains an item-specific channel capable of absorbing the additive shock generated by logarithmic single-good rescaling.

5.3 Hedonic regressions under imputation without commodity fixed effects

We now turn to hedonic imputation without commodity fixed effects. The key distinction from Proposition 5.1 is precisely the absence of such an item-specific absorption channel. Without it, a single-good rescaling generally affects the re-estimated model and therefore the imputed commodity-level price relatives themselves.

Consider a *pooled* hedonic regression estimated on an item–time (or commodity–country) sample Ω:

$$y_{ij} = f(P_{ij}) = \mu + \delta_j + \beta' z_{ij} + \varepsilon_{ij}, \qquad \sum_j \delta_j = 0, \tag{20}$$

where $f : \mathbb{R}_{++} \to \mathbb{R}$ is continuous and strictly increasing. We assume that the error term enters additively and is invariant under a single-good rescaling experiment: if the observed prices of one commodity are multiplied by a factor $K > 0$, the realizations ε_{ij} remain unchanged, so that the rescaling affects (20) only through the transformed dependent variable $y_{ij} = f(P_{ij})$. With fitted transformed value

$$\hat{m}_{ij}(P) = \hat{\mu}(P) + \hat{\delta}_j(P) + \hat{\beta}(P)' z_{ij},$$

and prediction rule

$$\widehat{P}_{ij}(P) = f^{-1}\big(\hat{m}_{ij}(P)\big),$$

the imputed commodity-level relatives are constructed from the resulting fitted or predicted prices.

In the model in (20), commensurability under imputation generally fails, even when $f(p) = a \ln p + b$. That is, there exist datasets P, commodities $i^\star$, and scale factors $K > 0$ such that

$$R_i\big(S_{i^\star,K}(P)\big) \neq R_i(P)$$

for some commodity $i \in I^\dagger$.

A simple illustration is provided by a two-commodity, two-column unbalanced panel in which commodity 1 is observed in both columns, commodity 2 is observed only in the base column, and the prediction equation is a log-hedonic specification without commodity fixed effects. Rescaling commodity 1 changes the fitted common component of the model and therefore changes the imputed price of commodity 2 in the comparison column. Hence the resulting relative for commodity 2 is not invariant.

This is the imputation counterpart of the failure of commensurability established in Section 4 for unbalanced hedonic models. Importantly, the source of the problem is not the choice of estimation method. It is not something peculiar to OLS, nor is it eliminated merely by replacing a linear hedonic regression with a more flexible predictor. The issue is instead whether the specification contains an item-specific channel capable of absorbing a single-good rescaling shock while leaving the resulting relatives unchanged.

For that reason, the same invariance criterion extends naturally to modern machine-learning- and AI-based hedonic imputation frameworks. Whatever is the estimation strategy, commensurability under imputation ultimately depends on the behavior of the predicted price surface under single-good rescaling.

Before turning to practice, note one narrow bridge case. In balanced panels with time-invariant characteristics and logarithmic transformation, no imputation is needed, since prices are observed in both periods for every commodity in every two-column comparison. In this special case, the present notion collapses to the earlier dummy-based notion. Thus the practically relevant imputation problem is that of the unbalanced case.

Definition 5.2 makes unit robustness under imputation precise: it requires invariance of all defined commodity-level relatives

$$\{R_i(P)\}_{i \in I^\dagger}.$$

to changes in units of measurement. The results above show that, under imputation, the basic conclusion is the same as in the earlier dummy-based analysis: what matters is whether the specification contains commodity-specific intercept shifts capable of absorbing the rescaling shock.

5.4 Imputation based on period-by-period regressions

The positive result in Proposition 5.1 relies on a *pooled* first-stage regression in which the commodity-specific intercept is common across periods (or countries). That common absorption channel is what allows a single-good rescaling under logarithmic transformation to be absorbed without changing the resulting commodity-level relatives.

This conclusion does *not* generally extend to imputation procedures that estimate a separate regression in each period. As emphasized by Silver and Heravi (2007), an important practical attraction of imputation-based index construction, relative to TPD or TDH approaches, is that period-by-period imputation allows the coefficients on observed attributes to vary over time and hence to adapt more flexibly to changing market conditions. However, this flexibility comes at the cost of commensurability.

In this subsection, we consider an imputation procedure that estimates a hedonic regression separately in each period. We focus on the case of one-sided missingness, that is, a commodity i for which, in the comparison pair $(t, 0)$, exactly one of (i, t) and $(i, 0)$ is observed. Under our missing-item imputation rule, in that case both sides of the prices are replaced by fitted values. We begin with the simplest intercept-only case.

Proposition 5.2 (Period-by-period log imputation with intercept-only regressions)**.** *Consider a two-period setting $t \in \{0, 1\}$. Let S_t denote the set of commodities observed in period t, and let $n_t := |S_t|$. Suppose that in each period one estimates a separate intercept-only regression*

$$\log P_{it} = \mu_t + \varepsilon_{it}, \qquad i \in S_t,$$

and uses the fitted value to impute one-sided missing prices. Then:

1. *The imputed price in period t is the geometric mean of observed prices:*

$$\widehat{P}_{it} = \exp(\hat{\mu}_t) = \Big(\prod_{k \in S_t} P_{kt}\Big)^{1/n_t} =: G_t.$$

2. *For any commodity i with exactly one-sided missingness across the two periods, the imputation-based relative is*

$$R_i(P) = \frac{\widehat{P}_{i1}(P)}{\widehat{P}_{i0}(P)} = \frac{G_1}{G_0}.$$

3. *Fix a commodity $i^\star$ and a scale factor $K > 0$, and apply the single-good rescaling*

$$P_{i^\star t} \mapsto K P_{i^\star t} \qquad \text{for all observed } t \in \{0, 1\}.$$

Then for every one-sided missing commodity i,

$$\frac{R_i\big(S_{i^\star,K}(P)\big)}{R_i(P)} = K^{d(i^\star)},$$

where

$$d(i^\star) := \frac{\mathbf{1}\{i^\star \in S_1\}}{n_1} - \frac{\mathbf{1}\{i^\star \in S_0\}}{n_0}.$$

4. *Let $B := S_0 \cap S_1$ be the set of commodities observed in both periods, let $M := S_0 \triangle S_1$ be the set of commodities with one-sided missingness, and let $N := |B| + |M|$ be the number of defined relatives entering the missing-item-only Jevons index. Recall that*

$$d(i^\star) := \frac{\mathbf{1}\{i^\star \in S_1\}}{n_1} - \frac{\mathbf{1}\{i^\star \in S_0\}}{n_0},$$

that is, the difference between the weight of the rescaled commodity $i^\star$ in the period-1 and period-0 geometric means. Then the Jevons index is

$$J(P) = \left(\prod_{i\in B} \frac{P_{i1}}{P_{i0}} \prod_{i\in M} \frac{G_1}{G_0}\right)^{1/N},$$

whereas after the single-good rescaling it becomes

$$J\big(S_{i^\star,K}(P)\big) = \left(\prod_{i\in B} \frac{P_{i1}}{P_{i0}} \prod_{i\in M} \frac{G_1'}{G_0'}\right)^{1/N}.$$

Hence

$$\frac{J\big(S_{i^\star,K}(P)\big)}{J(P)} = K^{\frac{|M|}{N} d(i^\star)}.$$

Therefore period-by-period intercept-only log imputation is generally not commensurable.

Proof. For each period t, OLS on the intercept-only regression gives

$$\hat{\mu}_t = \frac{1}{n_t} \sum_{k\in S_t} \log P_{kt},$$

so the fitted level prediction is

$$\widehat{P}_{it} = \exp(\hat{\mu}_t) = \exp\left(\frac{1}{n_t} \sum_{k\in S_t} \log P_{kt}\right) = \Big(\prod_{k\in S_t} P_{kt}\Big)^{1/n_t} =: G_t.$$

This proves part (i).

If commodity i has one-sided missingness, the missing-item-only imputation rule replaces both sides by predicted values, so

$$R_i(P) = \frac{\widehat{P}_{i1}(P)}{\widehat{P}_{i0}(P)} = \frac{G_1}{G_0},$$

which proves part (ii).

Now apply the single-good rescaling. In period t, the rescaled intercept becomes

$$\hat{\mu}'_t = \hat{\mu}_t + \frac{\mathbf{1}\{i^\star \in S_t\}}{n_t} \log K,$$

because the rescaling adds $\log K$ to exactly one observed transformed price whenever $i^\star \in S_t$. Therefore

$$G'_t = \exp(\hat{\mu}'_t) = G_t\, K^{\mathbf{1}\{i^\star \in S_t\}/n_t}.$$

Hence, for any one-sided missing commodity i,

$$R_i\big(S_{i^\star,K}(P)\big) = \frac{G'_1}{G'_0} = \frac{G_1}{G_0} K^{\mathbf{1}\{i^\star \in S_1\}/n_1 - \mathbf{1}\{i^\star \in S_0\}/n_0},$$

which proves part (iii).

Finally, the missing-item-only Jevons index based on all defined relatives is

$$J(P) = \left(\prod_{i \in B} \frac{P_{i1}}{P_{i0}} \prod_{i \in M} \frac{G_1}{G_0} \right)^{1/N}.$$

Under the rescaling experiment, the bilateral relatives for commodities in B are unchanged: if $i \neq i^\star$, neither observed price changes, and if $i = i^\star$, both observed prices are multiplied by the same factor K, so the relative cancels out. Only the M terms are affected, each by the common factor $K^{d(i^\star)}$. Thus

$$\frac{J\big(S_{i^\star,K}(P)\big)}{J(P)} = \Big(K^{d(i^\star)}\Big)^{|M|/N} = K^{\frac{|M|}{N} d(i^\star)},$$

which proves part (iv). □

Proposition 5.2 shows that the source of the failure is very simple. Under period-by-period intercept-only log imputation, the missing price in each period is replaced by that period's geometric mean of observed prices. A single-good rescaling therefore affects the imputed relative only through its effect on those period-specific means.

To make the result more transparent, write

$$\frac{J\big(S_{i^\star,K}(P)\big)}{J(P)} = K^{\rho\, d(i^\star)}, \qquad \rho := \frac{|M|}{N},$$

where

$$d(i^\star) = \frac{\mathbf{1}\{i^\star \in S_1\}}{n_1} - \frac{\mathbf{1}\{i^\star \in S_0\}}{n_0}.$$

Here $d(i^\star)$ is simply the difference between the weight that the rescaled commodity $i^\star$ receives in the period-1 geometric mean and the weight it receives in the period-0 geometric mean.

This gives three economically relevant cases. If $i^\star$ is observed in both periods, then

$$\frac{J'}{J} = K^{\rho(1/n_1 - 1/n_0)}.$$

If $i^\star$ is observed only in period 1, then

$$\frac{J'}{J} = K^{\rho/n_1}.$$

If $i^\star$ is observed only in period 0, then

$$\frac{J'}{J} = K^{-\rho/n_0}.$$

In particular, when $n_0 = n_1$ and the commodity is observed in both periods, the effect cancels out; otherwise the index generally changes.

These expressions make the size effect transparent. In particular, the deviation from commensurability is of order $1/n_t$: a single rescaled commodity affects a period-specific geometric mean only through its share in the sample for that period. Thus, holding the missing share ρ and the rescaling factor K fixed, the failure becomes smaller as the number of observed commodities in each period grows.

This also clarifies the practical relevance for scanner data. When n_t is large, as is often the case in scanner datasets, the bilateral deviation from commensurability may be mechanically small, so that J'/J can be close to one. However, Proposition 5.2 shows that the deviation is nevertheless generically nonzero, and even a small link-level distortion may accumulate under chaining.

Table 4 illustrates this in the simple case $K = 1000$, $\rho = 0.5$, and a rescaled commodity that is observed only in period 1. In that case,

$$\frac{J'}{J} = 1000^{\rho/n_1} = 1000^{0.5/n_1} = e^{(\ln 1000)\cdot 0.5/n_1} \approx 1 + \frac{0.5 \ln(1000)}{n_1}.$$

A further practical point is that even a small bilateral failure can accumulate under chaining. Suppose, for illustration, that the same stylized distortion recurs over T consecutive links, with the same sample size n in each link. Then, in the repeated-link case corresponding to Table 4, each bilateral link is distorted by

$$\frac{J'_{t-1,t}}{J_{t-1,t}} = 1000^{0.5/n},$$

Table 4: Magnitude of commensurability failure under period-by-period log imputation

n_1	Exact factor $1000^{0.5/n_1}$	First-order approximation $1 + 0.5\ln(1000)/n_1$
10	1.4125	1.3454
50	1.0715	1.0691
100	1.0351	1.0345
500	1.0069	1.0069

Notes. The table reports the multiplicative deviation J'/J from commensurability in the case $K = 1000$, $\rho = 0.5$, and $d(i^\star) = 1/n_1$, corresponding to a commodity $i^\star$ that is observed only in period 1. If commensurability held exactly, this ratio would be equal to one. The table shows that the deviation is positive but shrinks rapidly with the period-1 sample size: it is about 41.3% when $n_1 = 10$, about 7.2% when $n_1 = 50$, about 3.5% when $n_1 = 100$, and about 0.7% when $n_1 = 500$. Thus the failure is generally nonzero, but becomes mechanically small in large cross sections because one commodity has only a $1/n_1$-order influence on the period-specific geometric mean.

so the chained distortion becomes

$$\frac{C'_T}{C_T} = \prod_{t=1}^{T} \frac{J'_{t-1,t}}{J_{t-1,t}} = \left(1000^{0.5/n}\right)^T = 1000^{0.5T/n}.$$

Thus a per-link deviation that may appear moderate in a single bilateral comparison can become very large when distortions of the same sign recur across many links.

This numerical illustration is intentionally stylized and rests on a strong assumption: the link-level distortion is taken to have the same sign and the same magnitude in every link. In practice, the sign and size of the distortion may vary across links as sample composition, overlap patterns, and first-stage estimates change over time. Accordingly, actual chained drift need not be as large as in this repeated-link example.

Table 5 reports a numerical illustration for $n = 100$, so that each bilateral link is distorted by only

$$1000^{0.5/100} \approx 1.0351.$$

Even in this case, chaining causes the deviation to accumulate over time.

Section 7 provides an empirical illustration with scanner data. As will be seen there, the realized chained drift is substantially milder than in Table 5, in part because the sign of $d(i^\star)$ (the difference in its weights across periods) tends to alternate across links as products enter and exit the sample, and in part because each product's influence on the period-specific mean is of order $1/n_t$ in large cross-sections. Nevertheless, the chained drift remains economically non-negligible even in the full sample—reaching around two percent over the sample period—and is markedly larger in smaller stratified samples and under weekly imputation. This reinforces the practical relevance and implications of the failure of of the the commensurability axiom.

This non-commensurability is not confined to the intercept-only case considered above. It extends also to period-by-period regressions with observed attributes. Adding characteristics does not restore commensurability, because the essential problem is not the absence of regressors but the absence of a *common* commodity-specific absorption channel

Table 5: Accumulation of commensurability failure under chaining

Number of links T	Exact factor $1000^{0.005T}$	Percentage deviation
1	1.0351	3.51%
12	1.5136	51.36%
24	2.2909	129.09%
60	7.9433	694.33%
120	63.0957	6209.57%

Notes. The table assumes the same stylized link-level distortion in every link: $K = 1000$, $\rho = 0.5$, and $d(i^\star) = 1/n = 1/100$. Thus each bilateral link is distorted by the factor

$$1000^{0.5/100} \approx 1.0351,$$

and after T links the chained distortion is

$$\frac{C'_T}{C_T} = 1000^{0.5T/100} = 1000^{0.005T}.$$

The table shows that a bilateral deviation of about 3.5% per link can cumulate to about 51.4% after 12 links, 129.1% after 24 links, 694.3% after 60 links, and 6209.6% after 120 links. These magnitudes should be interpreted as a stylized repeated-link benchmark, not as a literal empirical prediction.

across periods. When the first-stage regression is re-estimated separately in each period, a single-good rescaling changes the fitted prediction function in the affected period, and therefore generally changes the imputed missing price and the resulting relative. Thus, even under logarithmic transformation, period-by-period hedonic imputation is generally not commensurable.

The results above are stated for parametric procedures whose fitted transformed values are well defined up to the common location normalization.

For more general prediction methods, the relevant requirement is that if the transformed prices of one commodity are shifted by a constant (as under logarithmic rescaling), the prediction rule should adjust only through a corresponding commodity-specific shift and should leave all bilateral relatives unchanged.

6 Numerical Examples

We now turn from analytical characterization to illustrative numerical examples. The purpose is not to search for new phenomena, but to show how the invariance to units of measurement results appear in a carefully constructed small and transparent unbalanced dataset.

The core theoretical results developed in Sections 3, 4, and 5 are summarized in Table 6. The numerical examples below operationalize these predictions on a small, deliberately unbalanced dataset.

The examples are organized in a sequence of three steps. We first describe the dataset, then define the single-good rescaling experiment, and finally report the consequences

Table 6: Summary of commensurability property in different models

		Log	**Box–Cox ($\lambda \neq 0$)**
Balanced	With Product Dummies	Yes	No
	With Attributes	Yes	No
	No Attributes	Yes	No
Unbalanced	With Product Dummies (pooled regression)	Yes	No
	Period-by-period imputation	No	No
	With Attributes (no product dummies)	No	No
	No Attributes	No	No

Notes. "Balanced" means that the same set of commodities is observed in every period (or country). "With Product Dummies" refers to specifications including commodity fixed effects in a pooled regression. "With Attributes" refers to hedonic specifications with observed characteristics but without commodity fixed effects, unless otherwise stated. The balanced hedonic "Yes" entry is to be understood under the time-invariant-characteristics assumption maintained throughout the paper. For the imputation results in Section 5, the positive logarithmic result with product dummies applies to a *pooled* regression in which commodity effects are common across periods (or countries). If instead the first-stage regression is estimated separately period by period, that common absorption channel disappears, and commensurability generally fails even under logarithmic transformation.

for two distinct objects: time-dummy-based bilateral comparisons and imputation-based price relatives.

6.1 Dataset

We construct a dataset with 4 commodities and 3 time periods (12 potential cells). The panel is unbalanced: Good 3 is missing in Period 2 and Good 4 is missing in Period 3. A continuous attribute z_1 is assigned by commodity: Good 1 ($z_1 = 0$), Good 2 ($z_1 = 1$), Good 3 ($z_1 = 3$), and Good 4 ($z_1 = 4$). The baseline prices are given in Table 7.

Table 7: Price matrix for each commodity–country pair

Commodity	**Period 1**	**Period 2**	**Period 3**
Good 1 ($z_1 = 0$)	1.0	1.1	1.2
Good 2 ($z_1 = 1$)	2.0	2.2	4.0
Good 3 ($z_1 = 3$)	3.0	—	8.0
Good 4 ($z_1 = 4$)	4.0	5.0	—

We consider the logarithmic transformation

$$f(p) = \ln p$$

and the Box–Cox family

$$f_\lambda(p) = \frac{p^\lambda - 1}{\lambda}$$

with $\lambda \in \{1, 0.5, 0.1\}$, where $\lambda = 1$ corresponds to levels. All transformations are applied directly to the observed prices; no mean normalization is performed prior to estimation.

6.2 Single-good rescaling experiment

We examine one primitive commensurability test: multiply all prices of Good 2 by $K = 10$, re-estimate each model, and compare the resulting bilateral comparisons and imputed price relatives.

This experiment is the numerical counterpart of the single-good rescaling as established in Proposition 5.1 in Section 3 . Under TPD commensurability, identified bilateral comparisons should be unchanged up to normalization. Under hedonic commensurability, the identified bilateral comparisons built from transformed levels should be unchanged. Under imputation commensurability, the commodity-level relatives used in the index should be invariant to this change in the measurement unit.

6.3 Time-dummy-based bilateral comparisons

We first report the implications of the rescaling experiment for bilateral comparisons recovered directly from the estimated models.

We estimate three specifications:

1. **TPD**: time and commodity fixed effects, with reference-country normalization $\delta_1 = 0$;
2. **Hedonic without attributes**: time dummies only;
3. **Hedonic with attributes**: time dummies plus a linear characteristic term $h(z_i) = \gamma z_{1i}$.

The TPD specification corresponds to the special case analyzed in Section 3, while the hedonic specifications correspond to the general time-dummy hedonic model and its unbalanced-panel implications in Sections 3 and 4. In the numerical implementation, all models are estimated in Stata using `cnsreg` with the normalization $\delta_1 = 0$.

In the logarithmic case, the reported bilateral comparison is

$$PPP(2,1) = \exp(\hat{\delta}_2 - \hat{\delta}_1) = \exp(\hat{\delta}_2) \qquad \text{under } \delta_1 = 0.$$

For Box–Cox specifications, the reported bilateral comparison is obtained by inverting the fitted transformed values and taking the corresponding level ratio. Thus, for a specification with fitted transformed value

$$\hat{m}_{ij} = \hat{\mu} + \hat{\delta}_j + \hat{\beta}' z_{ij},$$

the reported bilateral comparison is

$$\frac{\widehat{P}_{i2}}{\widehat{P}_{i1}} = \left(\frac{1 + \lambda \hat{m}_{i2}}{1 + \lambda \hat{m}_{i1}} \right)^{1/\lambda}.$$

In the TPD case, this becomes

$$\frac{\widehat{P}_{i2}}{\widehat{P}_{i1}} = \left(\frac{1 + \lambda(\hat{\mu} + \hat{\alpha}_i + \hat{\delta}_2)}{1 + \lambda(\hat{\mu} + \hat{\alpha}_i + \hat{\delta}_1)} \right)^{1/\lambda},$$

and under the normalization $\delta_1 = 0$,

$$\frac{\widehat{P}_{i2}}{\widehat{P}_{i1}} = \left(\frac{1 + \lambda(\hat{\mu} + \hat{\alpha}_i + \hat{\delta}_2)}{1 + \lambda(\hat{\mu} + \hat{\alpha}_i)} \right)^{1/\lambda}.$$

In the hedonic model with attributes, the corresponding expression is

$$\frac{\widehat{P}_{i2}}{\widehat{P}_{i1}} = \left(\frac{1 + \lambda(\hat{\mu} + \hat{\delta}_2 + \hat{\gamma} z_{1i})}{1 + \lambda(\hat{\mu} + \hat{\gamma} z_{1i})} \right)^{1/\lambda}.$$

Because Box–Cox reported bilateral comparisons depend on the evaluation point through the fitted linear predictor, the numerical example evaluates the hedonic-with-attributes case at the reference commodity with $z_{\text{ref}} = 0$ (Good 1). These reported PPPs are used only for numerical illustration; the theoretical notion of commensurability in the main text is defined directly in terms of identified bilateral comparisons or imputed relatives.

Table 8 reports the reported PPP for Country 2 relative to Country 1 under the baseline data and after the Good 2 rescaling experiment. Values ≥ 10 are rounded to two decimals; values < 10 to three decimals.

Table 8: Reported PPPs (Country 2 relative to Country 1)

Method	**TPD**		**Hedonic** ($h(z) = 0$)		**Hedonic** ($h(z) \neq 0$)	
	Base	**Rescale**	**Base**	**Rescale**	**Base**	**Rescale**
Log ($\lambda = 0$)	1.218	1.218	1.037	1.257	1.180	1.338
BC ($\lambda = 1$)	182.200	0.760	1.107	1.338	3.178	1.249
BC ($\lambda = 0.5$)	1.580	1.633	1.070	1.335	1.381	1.356
BC ($\lambda = 0.1$)	1.257	1.255	1.043	1.278	1.204	1.358

Notes. "Base" uses the original data. "Rescale" multiplies the prices of Good 2 by $K = 10$ and re-estimates the model. For the hedonic specification with $h(z) \neq 0$, the reported PPP under Box–Cox transformation is evaluated at the reference characteristic value $z_{\text{ref}} = 0$, corresponding to Good 1 in the numerical example. Thus, in the nonlinear case, the reported PPP is a predicted relative at a fixed reference characteristic value, rather than a comparison that is invariant to the evaluation point.

Interpretation (time-dummy-based PPPs). Table 8 illustrates the three main theoretical messages of the paper.

First, in the TPD model the reported PPP is exactly invariant under the logarithmic transformation ($1.218 \to 1.218$). This is exactly what the theory predicts: with $f = \log$, a single-good unit change is absorbed entirely by the corresponding commodity effect, leaving the identified bilateral comparison unchanged.

Second, TPD with Box–Cox transformations is not commensurable. The movement in the reported PPP (for example, 182.200 $\to$ 0.760 when $\lambda = 1$) illustrates the failure of the constant-difference property

$$f(Kp) - f(p) = C(K),$$

and hence the failure of commensurability outside the log case. The especially large base value for $\lambda = 1$ also shows that, under nonlinear transformations, the reported PPP can become highly sensitive to the level at which the inverse transformation is evaluated.

Third, the hedonic PPPs move even under logs on this unbalanced dataset (1.037 $\to$ 1.257 without attributes and 1.180 $\to$ 1.338 with attributes). This is consistent with Proposition 3.2 together with the results of Section 4: in hedonic models without commodity fixed effects, logarithms remain necessary, but they are generally *not sufficient* for commensurability on unbalanced panels.

Finally, when attributes are included, the Box–Cox hedonic PPP is a predicted relative evaluated at a reference characteristic value ($z_{\text{ref}} = 0$ here), so its numerical value depends not only on the estimated country effect but also on the level at which prediction is made. This is why the reported PPPs in the hedonic model with $h(z) \neq 0$ remain sensitive to the rescaling experiment, and why nonlinear transformations add an additional layer of evaluation-point dependence beyond the basic failure of commensurability.

6.4 Imputation-based price relatives

We next perform the same exercise under the imputation scheme described in Section 5. For any cell (i, c), observed or missing, define fitted prices by

$$\widehat{P}_{ic}^{\text{TPD}} = f^{-1}\big(\hat{\mu} + \hat{\alpha}_i + \hat{\delta}_c\big), \qquad \widehat{P}_{ic}^{\text{Hed}} = f^{-1}\big(\hat{\beta}_0 + \hat{\delta}_c + \hat{\beta}^\top z_i\big).$$

Under this rule, if both prices for commodity i in columns c and 1 are observed, we use the actual relative P_{ic}/P_{i1}. If exactly one of the two prices is missing, we replace both by fitted values and use $\widehat{P}_{ic}/\widehat{P}_{i1}$. Thus the primitive bilateral object is the collection of commodity-level relatives

$$\{R_i(P)\}_{i \in I^\dagger},$$

as in Definition 5.2.

Accordingly, commensurability under imputation requires invariance of these commodity-level relatives under the Good 2 rescaling experiment. Table 9 reports a summary for Country 2 relative to Country 1. For each specification, we construct the commodity-level relatives using observed prices when both prices are available and predicted prices only when exactly one side is missing, and then summarize the resulting set of relatives by a Jevons index. If the imputation procedure is commensurable in the sense of Definition 5.2, the ratio Rescale/Base must equal one.

Table 9: Imputation-based Jevons PPP (Country 2 relative to Country 1) under single-good scaling ($K = 10$)

Model	Base	Rescale	Ratio
TPD: log ($\lambda = 0$)	1.1651	1.1651	1.0000
TPD: Box–Cox ($\lambda = 1.0$)	1.1610	1.1976	1.0315
TPD: Box–Cox ($\lambda = 0.5$)	1.1639	1.1640	1.0001
TPD: Box–Cox ($\lambda = 0.1$)	1.1650	1.1646	0.9997
Hedonic with attributes: log ($\lambda = 0$)	1.1558	1.1928	1.0320
Hedonic with attributes: Box–Cox ($\lambda = 1.0$)	1.1549	1.1929	1.0329
Hedonic with attributes: Box–Cox ($\lambda = 0.5$)	1.1554	1.1933	1.0328
Hedonic with attributes: Box–Cox ($\lambda = 0.1$)	1.1557	1.1929	1.0322

Notes. "Base" uses the original data. "Rescale" multiplies the prices of Good 2 by $K = 10$ and re-estimates each model. Under the imputation rule used here, if both prices for a commodity are observed, the actual relative is used; if exactly one side is missing, both prices are replaced by their fitted values and the predicted relative is used. The reported Jevons PPP is therefore a summary of the commodity-level relatives $\{R_i(P)\}$ in Definition 5.2. "Ratio" is Rescale/Base. Under commensurability under imputation, Ratio must equal 1.

Interpretation (imputation-based Jevons PPP). Table 9 confirms the same qualitative pattern as the time-dummy-based comparisons, but now for the imputation object implied by Definition 5.2.

(i) **TPD + log** is exactly commensurable under imputation: the Jevons PPP is invariant to single-good rescaling (Ratio = 1). This is the numerical counterpart of Proposition 5.1: with $f = \log$, the rescaling shock is absorbed by the commodity effect, leaving all defined commodity-level relatives unchanged.

(ii) **TPD + Box–Cox** is not exactly commensurable under imputation, but the departures are now much smaller than under the earlier all-cells imputation exercise. This is because, under the present rule, observed relatives are retained whenever both prices are available, and prediction is used only for commodities with one-sided missingness. The resulting Jevons PPP therefore reflects only the limited part of the comparison that depends on imputation.

(iii) **Hedonic with attributes** fails commensurability under imputation even under logs in this unbalanced design: the Jevons PPP moves from 1.1558 to 1.1928. Thus, the instability is not confined to reported dummy comparisons, but extends to the commodity-level relatives used under imputation when the model lacks commodity fixed effects.

(iv) **Box-Cox tranformation** behave similarly under the present imputation rule. For all Box–Cox values considered, the Jevons PPP moves by about three percent under the single-good rescaling experiment, again indicating that the absence of commodity fixed effects, rather than the use of a particular nonlinear transformation alone, is the fundamental source of instability.

6.5 Discussion

Taken together, Tables 8 and 9 show that the same qualitative conclusion emerges from two distinct comparison strategies. Table 8 presents reported bilateral comparisons recovered directly from the estimated model, whereas Table 9 shows the Jevons indices computed using commodity-level relatives generated by the imputation rule. Thus, the two tables do not duplicate the same calculation; rather, they test commensurability under two different settings.

The numerical results are consistent across these two levels. In the TPD model with logarithmic transformation, single-good rescaling leaves both the reported bilateral comparison and the imputation-based relative structure unchanged. By contrast, outside the log-TPD case, instability appears in at least one of the two objects, and in the hedonic specifications on this unbalanced design it appears in both. In this sense, the imputation results reinforce, rather than replace, the message of the time-dummy-based results.

For applications that prioritize unit-rescaling invariance in both reported bilateral comparisons and imputed commodity-level relatives, log-TPD provides the natural benchmark. Hedonic approaches remain empirically informative, but in unbalanced data they require explicit robustness checks, because single-good rescalings generally propagate into the estimated bilateral comparisons and into the imputed relative structure.

7 Empirical Illustration with Scanner Data

To illustrate the empirical relevance of our theoretical results on commensurability, we analyze weekly scanner data for beverage products sold in Tokyo between January 2017 and December 2021. The data are drawn from the *SRI+* dataset compiled by Intage, Japan's leading marketing research company. We use weekly product-level data on sales and quantities aggregated over 104 supermarkets in Tokyo. The empirical application serves two purposes: first, to show that severe unbalancedness is a routine feature of scanner data; and second, to assess whether the unit-sensitivity identified in the theory is quantitatively important in practice.

We focus on five beverage categories: (1) 100% fruit juice, (2) fruit drinks (less than 100% juice), (3) cola, (4) cider, and (5) other carbonated beverages (excluding cola and cider).

Even after aggregation across stores, the number of product–week observations remains substantial:

A distinctive feature of scanner data is the frequent entry and exit of products. On average, each Japanese Article Number (JAN)[10] appears in the sample for only 55.8 out of 261 weeks. Only 4.5% of items (166 of 3,682) are present throughout the entire period, while 5.2% appear in only a single week. These figures underscore how unbalanced the

[10]The Japanese Article Number (JAN) is a 13-digit product code administered by GS1 Japan and is also known as GTIN-13.

Table 10: Distribution of product–week observations by beverage category

Category	**Freq.**	**Percent**
100% Fruit Juice	61,785	30.10%
Fruit Drinks (Not 100% Juice)	41,586	20.26%
Cola	13,717	6.68%
Cider	18,703	9.11%
Other Carbonated Beverages	69,485	33.85%
Total	205,276	100.00%

panel is, making the commensurability issues studied in Sections 3, 4, and 5 directly relevant.

Throughout this section, prices are measured as unit values on a unit-price basis, i.e. sales divided by physical quantity. To examine commensurability, we conduct a rescaling experiment in which only the 100% fruit juice category is converted from a per-milliliter basis to a per-liter basis. Equivalently, for that category only, quantities are divided by 1,000, so that unit prices and package-size variables are rescaled consistently.[11]

We compare four approaches. The first is a TPD(CPD) regression with product fixed effects and week dummies. The second is a time-dummy hedonic (TDH) regression with log package size as an attribute. The third and fourth are imputation-based Jevons indices, where missing one-sided relatives are completed using either a pooled hedonic regression with week dummies or separate weekly hedonic regressions. In all cases estimation is by weighted least squares using weekly sales shares as weights.

To examine sample-size dependence, we repeat the exercise on three samples: a 10% stratified sample of products, a 50% stratified sample, and the full sample. Stratification is by beverage category.

Figure 1 compares the official monthly CPI, converted to a weekly series by assigning each week the CPI of its calendar month, with the TPD and TDH indices in the full sample. Two points stand out. First, the TPD index is invariant to the rescaling experiment, so the baseline and rescaled series coincide exactly. This is the empirical counterpart of Proposition 3.3. Second, the TDH index is not invariant: the baseline and rescaled series diverge visibly, and the gap varies over time rather than remaining a constant proportional shift. Thus, in scanner data, the time-dummy hedonic index is directly sensitive to the arbitrary choice of measurement unit when the panel is unbalanced and product fixed effects are absent. This is the empirical manifestation of the theoretical failure established in Section 4.

To make the source of this instability more transparent, Figure 2 turns to the no-covariate time-dummy regression, that is, the specification with only a constant and time dummies. This figure is a level representation of the closed-form result in Corollary 4.2.1, combined with the baseline series itself. Let HB_t and HR_t denote the exponentiated time-dummy indices before and after rescaling the target category by $K = 1000$, under the normaliza-

[11] When the price unit is changed, we also rescale the capacity variable so that the hedonic specification remains dimensionally coherent.

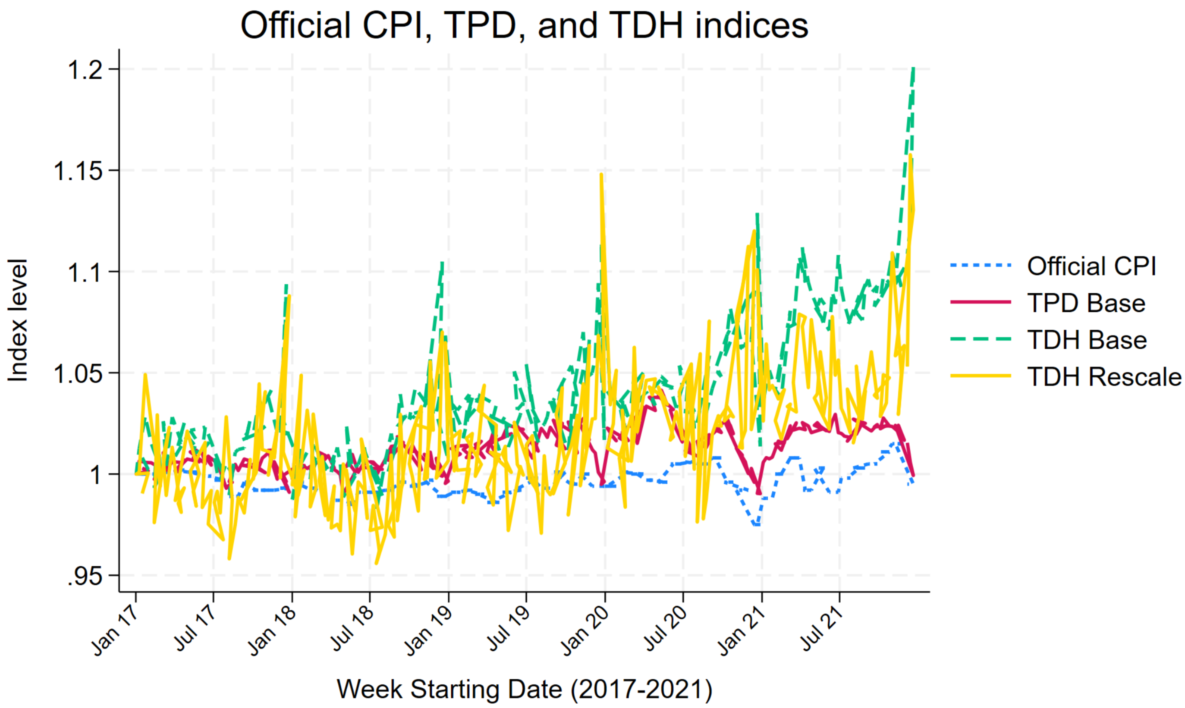

Figure 1: Official CPI, TPD, and TDH indices in the full sample. The TPD index is invariant to the change in the measurement unit, whereas the TDH index is not.

tion $\delta_1 = 0$.

Corollary 4.2.1 implies

$$\log HR_t - \log HB_t = \ln(K)(s_t - s_1),$$

where s_t is the weekly sales share of the rescaled category and s_1 is its base-week value. Hence, the divergence between HR_t and HB_t is generated entirely by time variation in the exposure of the rescaled category.

Figure 2 shows that the baseline no-covariate series HB_t does not differ dramatically from the baseline TDH series in Figure 1. After rescaling, however, the no-covariate series HR_t departs much more sharply from its baseline counterpart. Since TDH is commensurable on balanced hedonic panels by Proposition 3.4, this large divergence should be understood as arising from the unbalanced observation pattern rather than from the mere omission of attributes.

At the same time, comparison with Figure 1 shows that adding observed attributes such as ln(capacity) substantially attenuates this instability. In other words, observed characteristics mitigate the unit sensitivity generated by unbalancedness, but do not eliminate it completely. When the specification is enriched all the way to the TPD(CPD) special case with product fixed effects, commensurability is restored, as established in Proposition 3.3.

Next, we turn to the bilateral imputation results. Here, imputation is based on simple

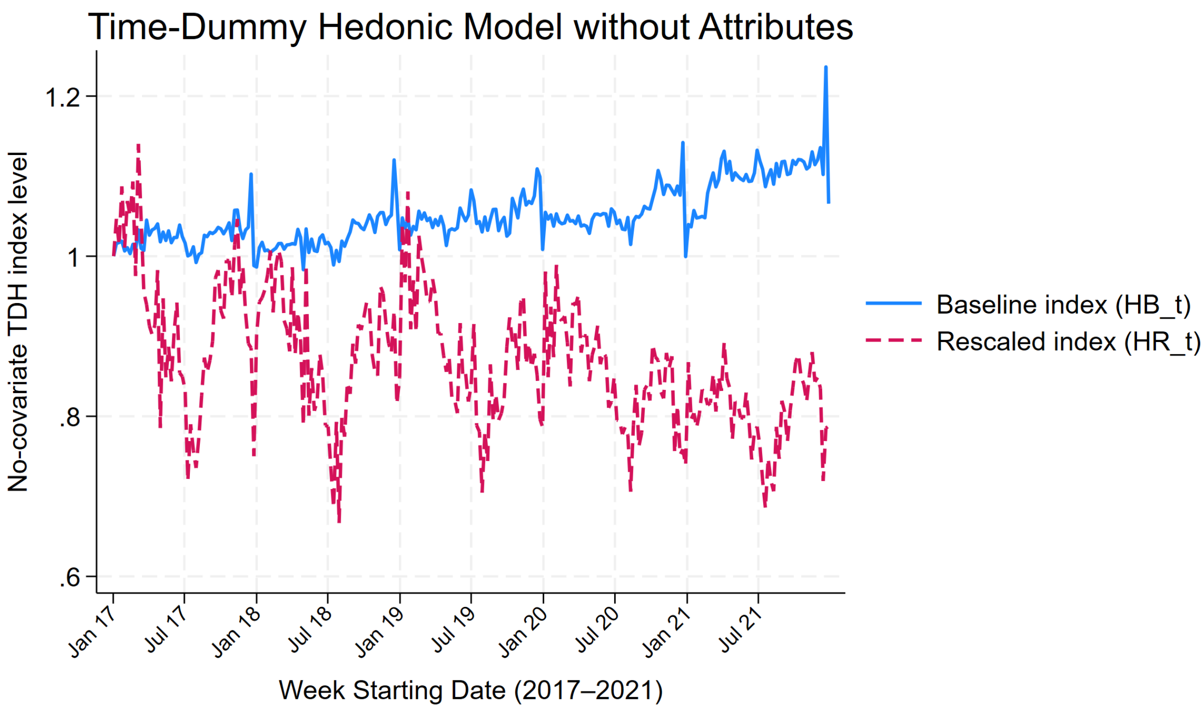

Figure 2: No-covariate time-dummy indices before and after rescaling the target category by $K = 1000$.

hedonic regressions in which log unit price is regressed on log capacity.[12] The pooled version is estimated on the full sample with week dummies,[13] whereas the weekly version is estimated separately for each week. Figure 3 shows that the effect of rescaling on *bilateral* imputation-based Jevons relatives is much smaller. For the full sample, both pooled and weekly imputation produce after/base ratios very close to one in each adjacent week pair. Hence, at the bilateral level, the direct effect of unit rescaling on imputation-based Jevons indices appears modest.

However, Figures 4 and 5 show that even small bilateral discrepancies can accumulate once the index is chained over time. For pooled imputation, the chained after/base ratio remains relatively close to unity in the full sample, but substantial drift emerges in the smaller samples, especially in the 10% stratified sample. For weekly imputation, the departures from unity are more pronounced still, and the divergence becomes persistent even for the 50% stratified sample. This contrast is intuitive: when the imputation regression is re-estimated separately each week, the effect of rescaling is allowed to vary more freely across adjacent links, and these link-specific distortions then accumulate through chaining.

[12] In the weekly specification, commodity dummies cannot be included, because with separate week-by-week estimation they are not identified.

[13] Using a pooled regression imposes a common prediction rule across weeks, but it does not make the resulting imputation-based bilateral Jevons indices transitive. Even under pooled estimation, the bilateral comparison remains based on commodity-level relatives that can differ across adjacent week pairs because the overlap pattern and the set of one-sided missing items vary over time.

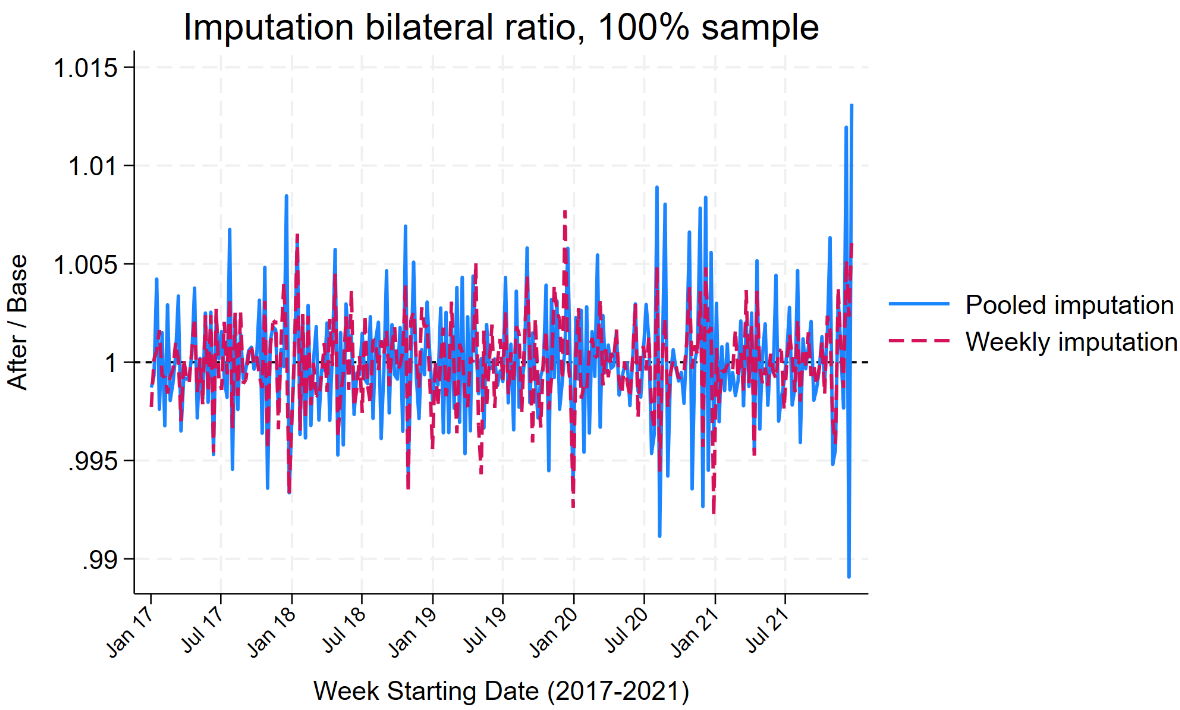

Figure 3: Bilateral imputation-based Jevons ratios in the full sample

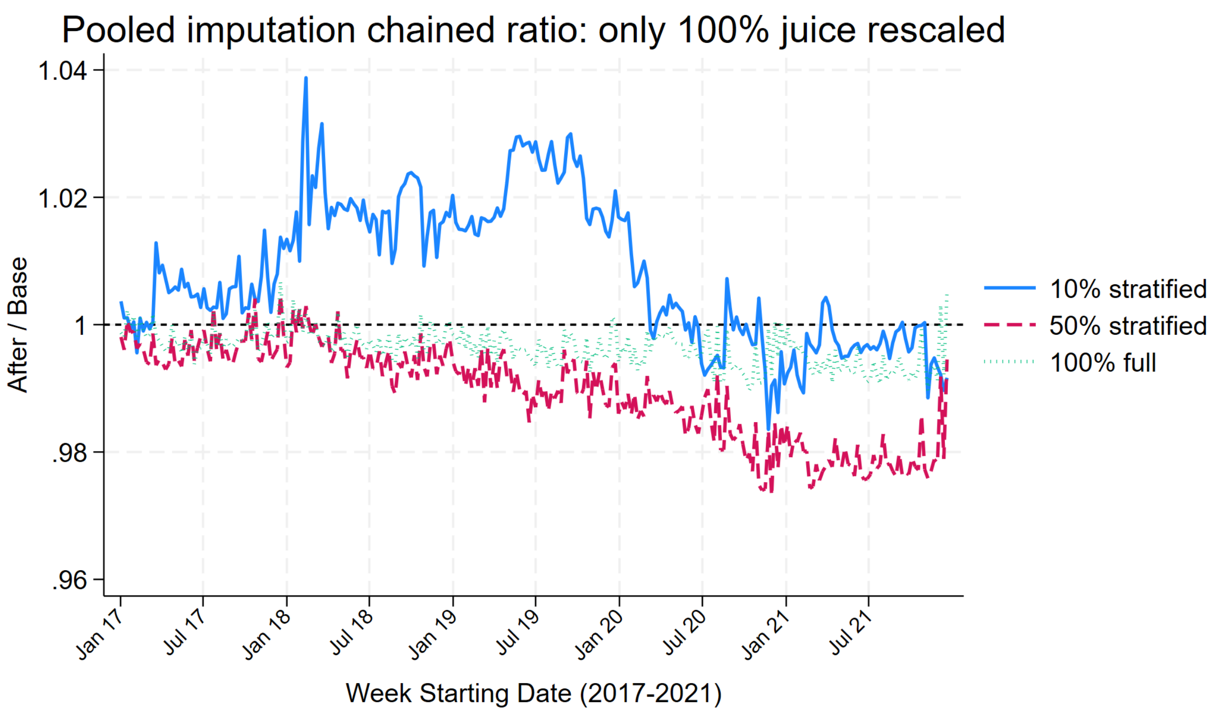

Figure 4: Chained commensurability ratio under pooled imputation

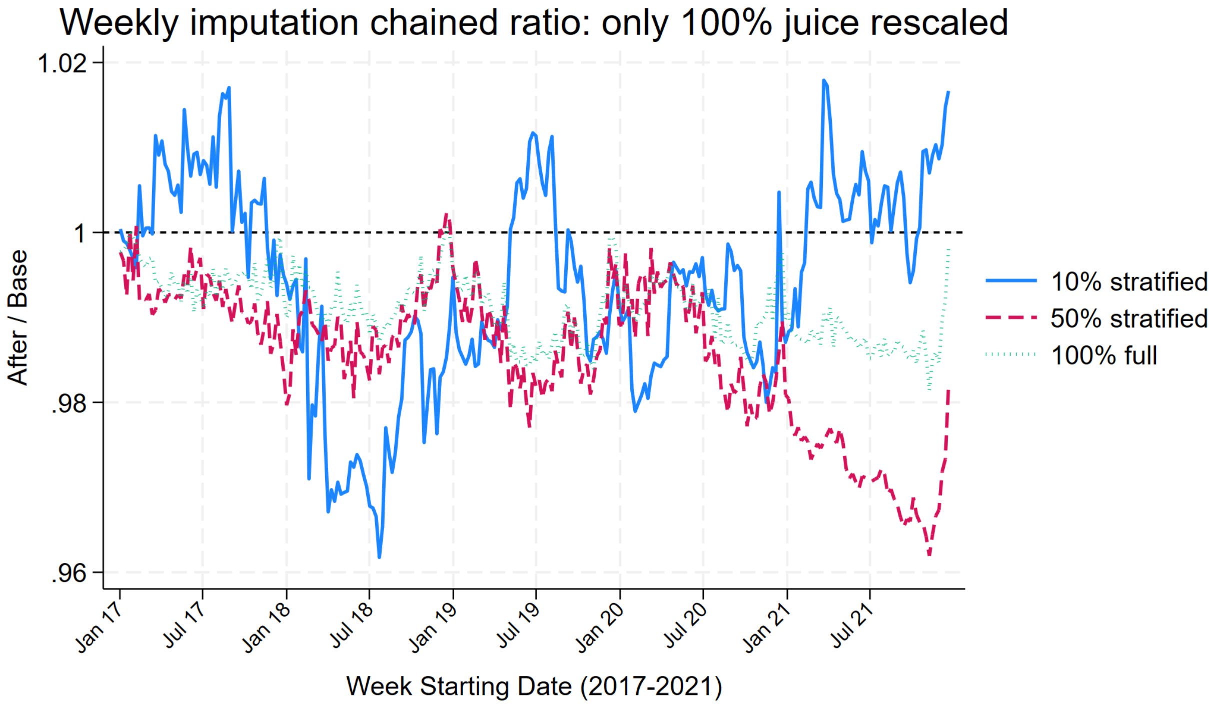

Figure 5: Chained commensurability ratio under weekly imputation

Figure 6 helps interpret this sample-size dependence. It plots the share of one-sided missing items in adjacent week pairs. This share fluctuates substantially over time and is markedly more volatile in the 10% stratified sample than in the larger samples. Because imputation affects only those relatives involving one-sided missingness, the practical importance of non-commensurability in imputation-based indices depends not only on the theoretical rescaling formula, but also on how many such missing links occur and how they are distributed over time.

Taken together, the scanner-data evidence supports four conclusions. First, the degree of unbalancedness observed in practice is large enough that commensurability is an empirically meaningful issue rather than a merely formal concern. Second, the TPD specification with product fixed effects remains commensurable under log transformation, as predicted by the theory, so its after/base ratio is omitted from the figures because it is identically one up to numerical error. Third, the TDH specification exhibits substantial non-commensurability already at the level of direct time comparisons, and the mechanism check shows that this is almost exactly explained by variation in the rescaled category's exposure over time. Fourth, imputation-based Jevons indices show much smaller bilateral effects, but even these small link-level departures can cumulate into noticeable chain drift, especially under weekly imputation and in smaller samples.

Overall, the scanner-data evidence is fully consistent with the theoretical message of Sections 3, 4, and 5: log transformations guarantee commensurability in dummy-based models with product fixed effects, but not in hedonic models without such fixed effects on unbalanced panels; and under imputation, even small link-level departures from com-

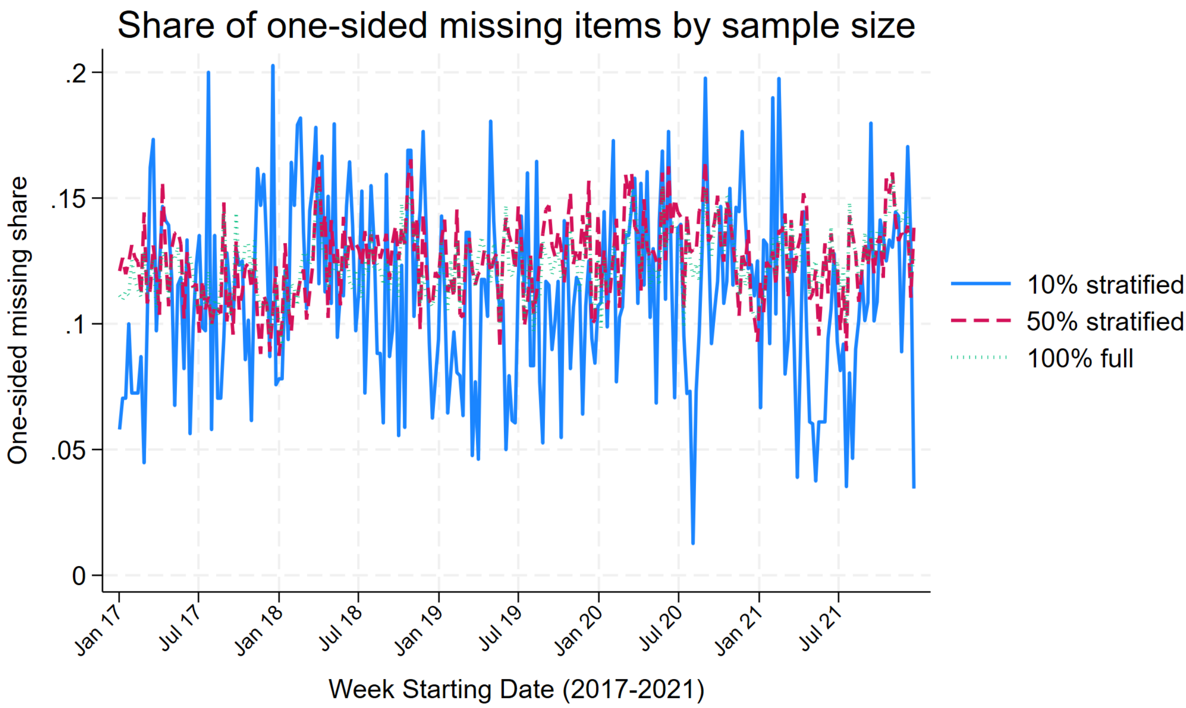

Figure 6: Share of one-sided missing items in adjacent week pairs

mensurability may cumulate into non-negligible chain drift.

8 Conclusion

This paper studies commensurability, namely invariance to changes in the measurement unit of a single commodity, in regression-based price index methods.

Our main theoretical result is that, in time-dummy models, the logarithmic transformation is not merely convenient but necessary once bilateral comparisons are constructed from normalization-invariant transformed levels and commensurability is imposed under single-item rescaling. In the TPD special case, the log transformation is also sufficient, because a single-good rescaling is absorbed by the corresponding commodity effect and leaves identified bilateral comparisons unchanged.

For time-dummy hedonic models, the conclusion is weaker but still important. The logarithmic transformation remains necessary for commensurability, but it is generally not sufficient in unbalanced panels. Without commodity fixed effects, the effect of a single-good rescaling typically propagates into the estimated bilateral comparisons through the projection structure of the model. Thus, with unbalanced data, hedonic methods are generally unit-sensitive even under logs. This implies that hedonic methods without commodity fixed effects should not be viewed as universally applicable procedures for broad and highly heterogeneous product spaces. Rather, they are most appropriately applied to relatively homogeneous product groups, where quality adjustment is central and the comparability of units can be more plausibly maintained.

The same distinction reappears under imputation, but in a practically important form. When the first-stage regression is pooled across periods and contains a common item-specific absorption channel, log-based specifications can deliver unit-invariant predicted relatives. By contrast, this result does not generally extend to period-by-period imputation. As emphasized by Silver and Heravi (2007), a major attraction of such procedures is that they allow the coefficients on observed attributes to vary over time and hence to adapt flexibly to changing market conditions. Our results show, however, that this flexibility comes at a cost: once the first-stage prediction rule is allowed to vary across periods, the common absorption channel disappears, and commensurability is generally lost even if commodity dummies are included in the period-specific regressions.

The empirical scanner-data results confirm that these distinctions are not merely formal. In Japanese beverage scanner data, the time-dummy hedonic index exhibits substantial departures from commensurability even at the level of direct time comparisons. Imputation-based Jevons indices display much smaller bilateral effects, but those small link-level discrepancies can still accumulate through chaining into noticeable drift, especially under weekly imputation and in smaller stratified samples. Thus, the practical importance of non-commensurability depends not only on functional form and model class, but also on how missingness is handled and whether comparisons are evaluated bilaterally or through chained aggregation.

The practical implications are straightforward. Unit invariance is not automatic when item prices are missing in some periods. If the main objective is a price index that is invariant to changes in units of measurement, log-TPD provides a natural benchmark. If the main objective is to track quality change, innovation, and time variation in attribute valuations, hedonic and imputation-based methods remain attractive, but their greater flexibility must be weighed against weaker robustness to arbitrary measurement conventions. Accordingly, without commodity fixed effects, such methods are best used for relatively homogeneous classes of goods, and unit harmonization together with explicit sensitivity checks should be treated as essential rather than optional. In this sense, applied price index construction involves a basic trade-off between axiomatic robustness and informational richness.

References

Aczél, János, *Lectures on Functional Equations and Their Applications*, Academic Press, 1966.

Asian Development Bank, *2017 International Comparison Program for Asia and the Pacific: Purchasing Power Parities and Real Expenditures — Results and Methodology*, Manila: Asian Development Bank, 2020.

Bajari, Patrick, Zhihao Cen, Victor Chernozhukov, Manoj Manukonda, Suhas Vijaykumar, Jin Wang, Ramon Huerta, Junbo Li, Ling Leng, George

Monokroussos, and Shan Wang, "Hedonic Prices and Quality Adjusted Price Indices Powered by AI," *Journal of Econometrics*, 2025, *251.*

Board of Governors of the Federal Reserve System, "Monetary Policy Report submitted to the Congress on June 20, 2025, pursuant to section 2B of the Federal Reserve Act," Technical Report, Board of Governors of the Federal Reserve System, Washington, DC June 2025.

Box, G. E. P. and D. R. Cox, "An analysis of transformations," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1964, *26* (2), 211–252.

Braun, Rahel and Sarah M. Lein, "Sources of Bias in Inflation Rates and Implications for Inflation Dynamics," *Journal of Money, Credit and Banking*, 2021, *53* (6), 1553–1572.

Broda, Christian and David E. Weinstein, "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, June 2010, *100* (3), 691–723.

Cafarella, Michael, Gabriel Ehrlich, Tian Gao, John C. Haltiwanger, Matthew D. Shapiro, and Laura Zhao, "Using Machine Learning to Construct Hedonic Price Indices," Technical Report Working Paper No. 31315, National Bureau of Economic Research June 2023.

Calainho, Filipe Dutra, "A Machine Learning Approach to Price Indices: Applications in Commercial Real Estate," *The Journal of Real Estate Finance and Economics*, 2024, *68* (4).

de Haan, Jan, "Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and 'Re-Pricing' Methods," *Jahrbücher für Nationalökonomie und Statistik*, 12 2010, *230* (6), 772–791.

Deaton, Angus, "Price Indexes, Inequality, and the Measurement of World Poverty," *American Economic Review*, March 2010, *100* (1), 5–34.

Diewert, W. Erwin, "Exact and Superlative Index Numbers," *Journal of Econometrics*, 1976, *4* (2), 115–145.

_ and Chihiro Shimizu, "Challenges in Consumer Price Index Construction," Discussion Paper 25-02, Vancouver School of Economics, The University of British Columbia, Vancouver 2025.

_ and _ , "Scanner Data, Product Churn and Quality Adjustment," NBER Working Papers 34897, National Bureau of Economic Research, Inc Feb 2026.

_ and Valentin Zelenyuk, "On the problem of the purchasing power of money by A. A. KonÃ¼s and S. S. Byushgens: translation and commentary," *Journal of Productivity Analysis*, April 2025, *63* (2), 107–132.

Erickson, Tim and Ariel Pakes, "An Experimental Component Index for the CPI: From Annual Computer Data to Monthly Data on Other Goods," *American Economic Review*, August 2011, *101* (5), 1707–38.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer, "The Next Generation of the Penn World Table," *American Economic Review*, October 2015, *105* (10), 3150–82.

Funke, Helmut, Günther Hacker, and Joachim Voeller, "Fisher's Circular Test Reconsidered," *Swiss Journal of Economics and Statistics (SJES)*, December 1979, *115* (4), 677–688.

Halvorsen, Robert and Henry O. Pollakowski, "Choice of functional form for hedonic price equations," *Journal of Urban Economics*, 1981, *10* (1), 37–49.

Handbury, Jessie, "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," *Econometrica*, 2021, *89* (6), 2679–2715.

Hill, Robert J., "Hedonic Price Indexes for Residential Housing: A Survey, Evaluation and Taxonomy," *Journal of Economic Surveys*, 2013, *27* (5), 879–914.

_ **and Iqbal A. Syed**, "Hedonic price-rent ratios, user cost, and departures from equilibrium in the housing market," *Regional Science and Urban Economics*, 2016, *56*, 60–72.

International Labour Office, "World Social Protection Report 2017–19: Universal Social Protection to Achieve the Sustainable Development Goals," Technical Report, ILO, Geneva 2017.

_ **, International Monetary Fund, Organisation for Economic Co-operation and Development, European Union, United Nations, and World Bank**, *Consumer Price Index Manual: Concepts and Methods*, International Monetary Fund, 2020.

Pakes, Ariel, "A reconsideration of hedonic price indexes with an application to PC's," *The American Economic Review*, 2003, *93* (5), 1578–1596.

Rao, D. S. Prasada, "The Framework for the International Comparison Program (ICP)," in World Bank, ed., *Measuring the Real Size of the World Economy*, Washington, DC: World Bank, 2013, chapter 1, pp. 13–45.

Ravallion, Martin, Gaurav Datt, and Dominique van de Walle, "QUANTIFYING ABSOLUTE POVERTY IN THE DEVELOPING WORLD," *Review of Income and Wealth*, 1991, *37* (4), 345–361.

Silver, Mick and Saeed Heravi, "The Difference between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes," *Journal of Business Economic Statistics*, 2007, *25* (2), 239–246.

Triplett, Jack, "Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products," Technical Report 2004/09, OECD Publishing, Paris 2004. OECD Science, Technology and Industry Working Papers, No. 2004/09.

A Functional Equations

This appendix collects the functional-equation results used in the main text. The following lemmas collect the functional-equation results used in the logarithmic necessity arguments. Lemma A.1 is used for the TPD-style constant-difference argument. Lemma A.4 is the projection-shift functional-equation result used in the time-dummy hedonic necessity proof, including both the OLS and WLS versions of the two-item CPD benchmark. Lemma A.3 records the final recovery step from the exponential form of ϕ to the logarithmic form of f. For classical treatments of such functional equations, see Aczél (1966).

A.1 Log-affine form from constant difference

Lemma A.1 (Log-affine form from constant difference)**.** *Let $f : \mathbb{R}_{++} \to \mathbb{R}$ be strictly increasing. Suppose that for every $K > 0$ there exists a scalar $C(K) \in \mathbb{R}$ such that*

$$f(Kp) - f(p) = C(K) \qquad \text{for all } p > 0. \tag{21}$$

Then there exist constants $a > 0$ and $b \in \mathbb{R}$ such that

$$f(p) = a \ln p + b \qquad \text{for all } p > 0.$$

Moreover,

$$C(K) = a \ln K \qquad \text{for all } K > 0.$$

Proof. Setting $p = 1$ in (21) gives

$$C(K) = f(K) - f(1).$$

Define

$$f_0(p) := f(p) - f(1).$$

Then f_0 is strictly increasing and satisfies

$$f_0(Kp) = f_0(K) + f_0(p) \qquad \text{for all } K, p > 0.$$

Now define

$$g(u) := f_0(e^u), \qquad u \in \mathbb{R}.$$

Because f_0 and $u \mapsto e^u$ are strictly increasing, g is strictly increasing. Moreover,

$$g(u + v) = f_0(e^{u+v}) = f_0(e^u e^v) = f_0(e^u) + f_0(e^v) = g(u) + g(v).$$

Thus g satisfies the additive Cauchy equation on $\mathbb{R}$ and is monotone. Hence there exists $a \in \mathbb{R}$ such that

$$g(u) = au \qquad \text{for all } u \in \mathbb{R}.$$

Therefore,

$$f_0(p) = g(\ln p) = a \ln p \qquad \text{for all } p > 0.$$

Since f is strictly increasing, we must have $a > 0$. Writing $b := f(1)$, we obtain

$$f(p) = a \ln p + b.$$

Substituting back into (21),

$$C(K) = f(K) - f(1) = a \ln K.$$

□

A.2 Exponential form from the shift equation

Lemma A.2 (Exponential form from the shift equation)**.** *Let $I \subseteq \mathbb{R}$ be an interval, and let $\phi : I \to \mathbb{R}_{++}$ be strictly increasing. Suppose there exists a set $J \subseteq \mathbb{R}$ containing a nondegenerate open interval such that, for every $v \in J$, there exists $\Psi(v) \in \mathbb{R}_{++}$ with*

$$\phi(x + v) = \Psi(v)\phi(x) \tag{22}$$

for all $x \in I$ such that $x, x + v \in I$. Then there exist constants $A > 0$ and $\lambda > 0$ such that

$$\phi(x) = Ae^{\lambda x} \qquad \text{for all } x \in I.$$

Proof. Because $\phi(x) > 0$ for all $x \in I$, define

$$\psi(x) := \ln \phi(x), \qquad x \in I.$$

Since ϕ is strictly increasing, ψ is strictly increasing as well. Taking logs in (22), we obtain

$$\psi(x + v) - \psi(x) = \ln \Psi(v) =: c(v) \tag{23}$$

for all $x \in I$ such that $x, x + v \in I$, and for all $v \in J$. Thus, for each admissible v, the increment $\psi(x + v) - \psi(x)$ depends only on v, not on x.

We claim that ψ is affine on I. Fix any open interval $U \subseteq J$. For $v, w \in U$ such that $v + w \in U$, we have

$$\psi(x + v + w) - \psi(x) = \big[\psi(x + v + w) - \psi(x + v)\big] + \big[\psi(x + v) - \psi(x)\big] = c(w) + c(v).$$

On the other hand, (23) also gives

$$\psi(x + v + w) - \psi(x) = c(v + w).$$

Hence

$$c(v + w) = c(v) + c(w)$$

whenever $v, w, v+w \in U$. So c satisfies the additive Cauchy equation on the open interval U. Because ψ is monotone, c is locally bounded, and therefore additive on U only if it is linear. Thus there exists $\lambda \in \mathbb{R}$ such that

$$c(v) = \lambda v \qquad \text{for all } v \in U.$$

Substituting back into (23),

$$\psi(x+v) - \psi(x) = \lambda v \qquad \text{for all admissible } x, v.$$

Fix $x_0 \in I$. For any $x \in I$, connect x_0 to x by finitely many increments lying in U. Summing the previous relation along that chain yields

$$\psi(x) - \psi(x_0) = \lambda(x - x_0).$$

Therefore

$$\psi(x) = \lambda x + b \qquad \text{for all } x \in I,$$

where $b := \psi(x_0) - \lambda x_0$. Exponentiating,

$$\phi(x) = e^{\psi(x)} = e^b e^{\lambda x} = A e^{\lambda x}, \qquad A := e^b > 0.$$

Finally, since ϕ is strictly increasing, we must have $\lambda > 0$. □

A.3 Recovering the logarithmic transformation

The previous lemma identifies the recovery map ϕ as exponential. The next observation records the final step used repeatedly in the main text.

Lemma A.3 (Recovering the logarithmic transformation). *Suppose $\phi : I \to \mathbb{R}_{++}$ is given by*

$$\phi(x) = A e^{\lambda x}, \qquad A > 0,\ \lambda > 0,$$

and suppose ϕ is the recovery map associated with the price transformation f, so that

$$\phi(f(p)) = p \qquad \textit{for all } p > 0.$$

Then

$$f(p) = \frac{1}{\lambda} \ln p - \frac{\ln A}{\lambda} = a \ln p + b$$

for some $a > 0$ and $b \in \mathbb{R}$.

Proof. From $\phi(f(p)) = p$ and $\phi(x) = A e^{\lambda x}$,

$$p = A e^{\lambda f(p)}.$$

Taking logs,

$$\ln p = \ln A + \lambda f(p).$$

Rearranging,

$$f(p) = \frac{1}{\lambda}\ln p - \frac{\ln A}{\lambda}.$$

Since $\lambda > 0$, the coefficient on $\ln p$ is positive. $\square$

A.4 Functional equations for logarithmic necessity

Lemma A.4 (Shift-ratio functional equation)**.** *Let $\varphi : I \to \mathbb{R}_{++}$ be strictly increasing and C^1, with $\varphi'(x) > 0$, and let $f = \varphi^{-1}$. Suppose that, for every $K = e^\kappa > 0$ and every admissible α, the function*

$$H_{K,\alpha}(x) := \frac{\varphi\big(x + \frac{1}{2}\{f(K\varphi(x+\alpha)) - (x+\alpha)\}\big)}{\varphi(x)}$$

is independent of x. Then

$$\varphi(x) = Ae^{\lambda x}$$

for some $A > 0$ and $\lambda > 0$. Equivalently,

$$f(p) = a \ln p + b, \qquad a > 0.$$

Proof. Write

$$L(x) := \ln \varphi(x), \qquad h(x) := L'(x) = \frac{\varphi'(x)}{\varphi(x)}.$$

Since $\varphi'(x) > 0$ and $\varphi(x) > 0$, we have $h(x) > 0$.

The assumption says that

$$L\bigg(x + \frac{1}{2}\{f(K\varphi(x+\alpha)) - (x+\alpha)\}\bigg) - L(x)$$

is independent of x.

Let $K = e^\kappa$. Define

$$D_\kappa(u) := f(e^\kappa \varphi(u)) - u.$$

Then the condition becomes

$$L\bigg(x + \frac{1}{2}D_\kappa(x+\alpha)\bigg) - L(x) = C(\kappa, \alpha),$$

where $C(\kappa, \alpha)$ is independent of x.

At $\kappa = 0$, we have $D_0(u) = 0$. Differentiating the preceding equation with respect to κ at $\kappa = 0$, we obtain

$$\frac{1}{2}h(x) \left.\frac{\partial D_\kappa(x+\alpha)}{\partial \kappa}\right|_{\kappa=0} = C_\kappa(0, \alpha).$$

Now

$$D_\kappa(u) = f(e^\kappa \varphi(u)) - u,$$

so

$$\left.\frac{\partial D_\kappa(u)}{\partial \kappa}\right|_{\kappa=0} = f'(\varphi(u))\varphi(u).$$

Since $f = \varphi^{-1}$,

$$f'(\varphi(u)) = \frac{1}{\varphi'(u)}.$$

Therefore

$$\left.\frac{\partial D_\kappa(u)}{\partial \kappa}\right|_{\kappa=0} = \frac{\varphi(u)}{\varphi'(u)} = \frac{1}{h(u)}.$$

Hence

$$\frac{1}{2}\frac{h(x)}{h(x+\alpha)} = C_\kappa(0,\alpha).$$

Thus, for each admissible α,

$$\frac{h(x+\alpha)}{h(x)}$$

is independent of x. Therefore there exists a positive function $A(\alpha)$ such that

$$h(x+\alpha) = A(\alpha)h(x).$$

Since h is continuous and positive, this multiplicative translation equation implies

$$h(x) = He^{\rho x}$$

for some $H > 0$ and $\rho \in \mathbb{R}$.

It remains to show that $\rho = 0$. Suppose, to the contrary, that $\rho \neq 0$. Then

$$L'(x) = He^{\rho x},$$

so

$$L(x) = B + \frac{H}{\rho}e^{\rho x}.$$

Thus

$$\varphi(x) = \exp\left(B + \frac{H}{\rho}e^{\rho x}\right).$$

Now return to the original finite-κ condition. Since

$$L(f(e^\kappa \varphi(u))) = L(u) + \kappa,$$

and

$$L(u) = B + \frac{H}{\rho}e^{\rho u},$$

we obtain

$$e^{\rho f(e^{\kappa}\varphi(u))} = e^{\rho u} + \frac{\rho}{H}\kappa.$$

Therefore

$$D_\kappa(u) = f(e^{\kappa}\varphi(u)) - u = \frac{1}{\rho}\ln\left(1 + \frac{\rho\kappa}{H}e^{-\rho u}\right).$$

Substituting $u = x + \alpha$, the left-hand side of the finite-κ condition is

$$L\left(x + \frac{1}{2}D_\kappa(x+\alpha)\right) - L(x).$$

Using $L(x) = B + \frac{H}{\rho}e^{\rho x}$, this becomes

$$\frac{H}{\rho}e^{\rho x}\left[\left(1 + \frac{\rho\kappa}{H}e^{-\rho(x+\alpha)}\right)^{1/2} - 1\right].$$

For $\kappa \neq 0$, this expression depends on x whenever $\rho \neq 0$. This contradicts the assumption that the expression is independent of x. Hence $\rho = 0$.

Therefore

$$h(x) = H$$

is constant. Since $h = L'$, we have

$$L(x) = B + Hx.$$

Thus

$$\varphi(x) = e^{B}e^{Hx} = Ae^{Hx}.$$

Because φ is strictly increasing, $H > 0$. Hence

$$\varphi(x) = Ae^{\lambda x}, \qquad \lambda > 0.$$

Since $f = \varphi^{-1}$,

$$f(p) = \frac{1}{\lambda}\ln(p/A) = a\ln p + b, \qquad a > 0.$$

This proves the lemma. □