Hedonic price estimates for new vehicles: When do rotations lead to drift?

Brendan K. Williams

May 2023

WORKING DRAFT

Prepared for the Group of Experts on Consumer Price Indices

UNECE, Geneva, June 2023

Abstract

Using a transaction level dataset on car purchases, we document the empirical relationship between standard hedonic index methods, including hedonic imputation and time dummy estimates, and the matched-model approach. We extend this analysis to investigate the effects of product cycles on hedonic estimates and the potential for coefficient "drift" that may result in biased price indexes. We distinguish between these effects and conventional "chain drift" in the context of bilateral, pooled, similarity linking, and multilateral index approaches. We map transaction records to additional sources to incorporate more detailed vehicle attributes and performance metrics. We introduce a new method of similarity linking specific to hedonic regression. Our results offer guidance on hedonic index construction methods and incorporating alternative data for industries into official price statistics.

JEL Codes: C43, E31

Brendan K. Williams is a Senior Economist in the Branch of Consumer Prices at the Bureau of Labor Statistics (BLS).
This research arose out of an earlier project with Leonard Nakamura, Ryan Michaels, and Erik Sager. Thank you to Bill Thompson and Nicole Shepler for their comments and review.

1. Introduction

The availability of transaction and scanner data has created a shift in focus on price index research, with methods related to chain drift receiving a great deal of attention. In many cases, downward drift may be generated by a product cycle and not the conventional "chain drift" mechanism, product cycle effects have received relatively little attention in the literature. We conduct our empirical analysis on new vehicle sales data, which have documented product cycle effects related to intertemporal price discrimination and price changes being introduced simultaneously with product updates. In a traditional, fixed-sample consumer price index, item replacement is used to address issues that result from this drift and quality adjustment emerges as a necessity to address the quality bias that may ensue. Hedonic methods are associated with measuring technological improvements and have negative price index impacts as a result. However, when hedonic imputation methods are used on data with a product cycle, we often see positive effects when hedonic imputation allows improved measurement of long-run price change.

The relationships between product cycles and quality change, and how they relate to various price index construction methods have not been well investigated. We use a dataset of new vehicle sales records to compare several, standard approaches to price index construction to analyze the empirical effects of product cycles. New vehicles were the subject of the very first hedonic analysis and have continued to be a subject of interest in the literature. We revisit some earlier hedonic models of vehicles and combine these specifications with more recent hedonic index methods.

When using transaction data product matching and grouping or hedonic imputation may be used to address drift. Multilateral methods play a role in offsetting product cycle drift, but mainly as a means of time aggregating hedonic imputations and addressing drift they may induce. We find that multilateral methods applied without hedonic imputation or product matching do not address product cycle drift. Within just the past few years, similarity linking methods have shown promise for dealing with chain drift as an alternative to GEKS-type multilaterals (Diewert, 2021). In an apparent first, we combine similarity linking methods with hedonic imputation. Similarity linking methods entail finding similar time periods for bilateral price index comparisons. Several different methods have been proposed to quantify "similarity" in practice. We introduce a new method where Chow test statistics are used with hedonic regression estimates to assess relative similarity between time periods.

We begin with an explanation of our data. We receive transaction level data from J.D. Power. These data are currently in estimation for the U.S. CPI based on the methodology in Williams and Sager (2019). We match these sales observations to more detailed specification information, including measures of vehicle performance, from Wards. The addition of this information allows us to produce more detailed hedonic models and reproduce specifications of historical interest.

We move to a discussion of previous work on hedonic estimates for vehicles. We conduct our empirical analysis on new vehicle sales. New vehicles have a long history in hedonic research going back to the seminal papers introducing and popularizing hedonic methods. We revisit these earlier model specifications and evaluate them in terms of more recent hedonic price index methods.

Next, we review evidence for product cycles and the intuition behind their effect on price indexes.

We then review the existing, established methodologies for using hedonic, multilateral, and similarity linking methods. We detail the use of our novel similarity linking method based on hedonic imputation.

Much of this paper is devoted to documenting the behavior of various standard hedonic and other price index methods on a single data source with well-documented product cycle behavior. We show that matched model price indexes tend to show implausibly large decreases. Hedonic estimates tend to show less of a decline. Our results are consistent with several other papers where matched model indexes are downwardly biased. We find that hedonic imputation methods may address product cycle issues by allowing long-run price comparisons to be made over a long time horizon.

2. Data

Transaction records with pricing data from J.D. Power were linked to specification data from Wards. The two sources were not consistent with each other, especially when identifying trim and packages. An algorithmic-assisted process was used to link records between the two sources. Concatenations of the model, trim, and package fields for both sources were created and compared. One-to-one matches were assumed to be correct. When a J.D. Power record matched several potential Wards vehicles, records were prioritized based on the highest number of words in common followed by the minimum string distance. These matches were then manually reviewed. The Wards data did not cover all observations in the J.D. Power data with certain trims and even a few models were omitted.

Our Wards data contain specification information from the 2005 to 2019 model year. 2020 model year vehicles began sales in 2019, meaning our specification information only covers our transaction data through 2018. Data for the 2019 model year also does not include mileage estimates (presumably these were not available at the time when we received this data from Wards), which limits model specifications that include fuel efficiency to the index through 2017.

This paper will focus on the pricing of passenger cars (defined as vehicles with listed body types of sedans, convertibles, hatchbacks, coupes, and wagons). Truck and van vehicle configurations (such as cabin type, bed length, and van height) can add variation to price, and our data does not indicate these specifications consistently. Moreover, many of these configurations are intended for commercial use and outside the scope of a Consumer Price Index.

Unlike much of the research published by BLS, the indexes created here are not intended as candidates for production use in the U.S. CPI. We have made many simplifications that would not be used in a production series (such as estimating a single-stage price index at the national level rather estimating area level indexes and performing an aggregation).

3. Background on New Vehicle Hedonic Research

New vehicle pricing has been the focus of landmark hedonic[1] methods papers including the foundational research in Court (1939) and the popularization of hedonics following Griliches (1961)—and more recently in the broader demand estimation literature with Berry, Levinsohn, and Pakes (1995). These early hedonic papers focused on tangible aspects of vehicles and their performance with Court

[1] Court first used the term “hedonic” and attributed the name to a suggestion from Alexander Sachs. Court’s paper is typically described as the beginning of hedonic, and it seems to have laid theoretical foundations, but other papers in agriculture proceeded it in using features as predictors of price. See Colwell and Dilmore (1999).

proposing a three-variable specification of weight, wheelbase, and horsepower. Griliches (1961) added dummy variables for V8 engines, hardtops, transmission, compact body type, and power brakes and steering. Triplett (1969) followed this specification but combined power brakes and steering and also proposed a truncated model. Cowling and Cubbin (1972) introduced several other variables including vehicle fuel efficiency.

Of these historical specifications, we can most directly reproduce the specifications used in Court (1939) and Ohta and Griliches (1976). Power brakes have long been standard on almost all vehicles sold in the United States and only a few models are available with manual steering (even manual transmission has become uncommon). The simple, three-variable model in Court remains directly applicable even to modern vehicles. The other models are less applicable as options like power brakes and steering are now nearly universally standard and while others no longer exist—namely, the pillarless "hardtop," which has not been sold since the 1970s. Omitting hardtop as an obsolete feature, the Ohta and Griliches specification is producible given our data and has the advantage of accounting for vehicle make. Since "make" is generally indicative of the level finishings in a vehicle, including make gives a rough control on interior quality (an aspect generally otherwise omitted in our data and the papers discussed below).

Table 1: Comparison of historical model specifications

	Court (1939)	Griliches (1961)	Triplett (1969)	Triplett Trunc. (1969)	Cowling & Cubbin (1972)	Ohta & Griliches (1976)
Weight	x	x	x	x		x
Wheelbase	x	length/wheelbase				
Horsepower	x	x	x		x	x
Length		length/wheelbase	x		x	x
V8		x	x			x
Hardtop		x	x			x
Transmission		x	X	Comb.		
Power brakes		x	Comb.	Comb.	x	
Power steering		x	Comb.	Comb.		
Compact		x	x	x		
Over4Gears					x	
Luxury					x	
PassengerArea					x	
Efficiency					x	
Make						Indicator variables

4. Product Cycle

We refer to regular patterns in product entry and exit and their effects price and quantity measurement as "product cycles." For cars, we focus on two elements of the product life cycle: price declines over a single product iteration driven by intertemporal price discrimination and the tendency for price change to be associated with model updates. Both of these elements of the vehicle product cycle lead to potential bias in estimating price change. Taking a simple matched model approach with product entry and exit through overlap would result in persistent downward index movement since price discrimination leads to a strong tendency for price discounts over a single iteration. Similarly, if sellers update their pricing strategies with new product entry, a matched model index with overlap would not reflect the change between pricing regimes. While these product cycle effects pertain to the new vehicle industry, other item categories also exhibit product cycle behavior that result in similar measurement issues.

Aizcorbe, et al. (2010) and Williams and Sager (2019) document evidence for intertemporal price discrimination related to consumer heterogeneity over the product cycle. Chained price comparisons that reflect the price change across variants fail to offset product life cycle effects. Williams and Sager (2019) found multilateral indexes without linking across product versions failed to counter downward drift and proposed a year-over-year, model-on-model measurement for the trend price in order to avoid the effects of price discrimination and account for price change with product updates.

Reinsdorf, et al. (1996) noted that sellers often introduced prices alongside new models. If indexes only show price change for the same version of an item (and overlap old and new products as they enter and exit the market), price change between regimes, which is the most important in the long run, will be omitted. In cases when sellers update their price strategy or schedules for inflation at the time of changing product offerings, omitting this price change will result in downward bias. This issue is addressed in traditional fixed sample surveys by showing price change between new and replacement items. Williams (2021) finds that the effects of item replacement and class-mean imputation, which is motivated by the need to capture and impute price change across products updates, are very large—larger than estimated quality bias in the index. Moreover, the need to correct for quality bias during these comparisons is ultimately motivated by the need to measure the price change across updates.

In a scanner data context, product matching and grouping have often been used. However, results can be sensitive to the producers used to map products together. Moreover, the timing and other aspects of the item replacement and dynamic weighting further complicate translating "item replacement" methods to scanner data.

Another approach is to use hedonic estimation. Looking at apparel data, Greenlees and McClelland (2010) found a similar pattern that we see in the vehicle market where prices decline strongly in within version price change. Their results varied greatly depending on the specific technique of hedonic index construction used. Multilateral methods did not address product cycle effects as "the relentless downward march of prices completely overwhelm the chain drift issue." Below we investigate various approaches to hedonic index construction and how they relate to product cycles.

5. Hedonic Methods

Hedonic methods basically predict a product's price as a function of its attributes. We can revisit the early model in Court to serve as an example.

$$Ln(Price) = \alpha + \beta_1 \times Wheelbase + \beta_2 \times Weight + \beta_3 \times Horsepower$$

Using linear regression, Court estimated the coefficient values (for wheelbase in inches, weight in hundredweight) in a joint time period regression for 1925 to 1930:

$$Ln(Price) = 4.1256 + 0.0161 \times Wheelbase + 0.0461 \times Weight + -0.0003 \times Horsepower$$

To estimate the price of a Model T in 1925 we can enter in the specification values for the Model T:[2]

$$Ln(Price) = 4.1256 + 0.0161 \times 100 + 0.0461 \times 12 + -0.0003 \times 20$$

This estimates the price of a Model as $535.29 in 1925.

We can perform the same exercise using a model estimated on modern data and estimate what a new Model T would cost in 2019 as

$$Ln(Price) = 9.9 + -0.0149 \times 100 + 0.0004 \times 1200 + 0.0029 \times 20$$

an estimate of $7673.06.

Hedonic methods are often suggested to address selection bias related to the immediate entry and exit of a product. While the literature generally expects hedonic indexes or hedonic adjustment to have downward impact on indexes, BLS research has found that hedonic adjustment has small or even upward effects shows that, were the BLS to omit the item replacement process entirely, indexes would generally be substantially lower. Previous research has found that these adjustments of have little impact on the U.S. CPI (Brown and Stockburger, 2006; Johnson, et al., 2006; Williams, 2021). Williams (2021) finds that product cycle effects are much larger than estimates of quality bias. Here, we focus on hedonic imputation as a means of calculating long-run price change in order to address these product cycle effects.

Model

We continue in the vein of the previous model discussed above in focusing on vehicle performance attributes and basic elements of vehicle size. In addition to the horsepower, we also have data on torque, and mileage broken out into city and highway estimates. Automotive engineers face basic tradeoffs in terms of power, weight, and efficiency. We create a highly interacted model to allow parameter estimates to account for the underlying relationships between these variables and better fit our data. We produce the Court and Ohta and Griliches models for their ease of interpretation and historical interest. The Ohta and Griliches specification is used as a benchmark for several comparisons in this paper.

[2] Court originally estimated with weight as "hundred weight," so we use 12 instead of 1200.

Hedonic Imputation Indexes

Hedonic imputation (HI) is often cited as the preferred approach to hedonic indexes (Diewert, 2019). In a hedonic imputation index, a hedonic regression is estimated on each period. The prices for the sets of goods in other periods are then estimated. Following the example above, taking a Model T imputed price from 1925, $535.29, and the imputed price of the Model T in 2019 gives us a Laspeyres price index increase of 1422% which is not far off from the overall CPI change of 1455% for the same period.

Silver and Heravi (2007) find them preferable to adjacent period time dummy hedonics (TDH). Unlike time product dummy (TPD) and TDH indexes, the dependent variable of price is not restricted to the natural log transformation. In TPD and TDH approaches, the dependent variable must be in the form of a natural logarithm to allow the time dummy to be interpreted as a proportional change in price.

Here we focus on full imputation, HI indexes where both omitted and observed are replaced with the predicted value produced by a hedonic regression for the corresponding period. Imputed "missing" observations comprise of an imputed price with a zero-value quantity and expenditure weight.

In our approach, all observations with the same set of values for a given specification are grouped together into one unit. For a detailed specification, this is equivalent or nearly equivalent to defining a unit by product identifier. In less detailed specifications, such as the Court model, a unit consists of transactions from multiple product identifiers.

Time Dummy Hedonic

Time dummy hedonic regressions constrain coefficients to have the same value over time. If the underlying parameter shifts between periods, the residual will be correlated with time period. The time dummy variable will then capture the difference.

In a time dummy model estimated on pooled dataset, the coefficient data is pooled over a long-time period. Another approach is to use an adjacent period TDH where a series of regressions is estimated on data pairs of adjacent periods with a dummy variable indicating the later period. These time dummy variables can be accumulated into a chained multiperiod index. The adjacent period index allows the

The time dummy hedonic equation is

$$\ln p_i^t = \alpha + \sum_{t=1}^{T} \delta^t D^t + \sum_{k=1}^{K} \beta_k z_{ik}$$

Given the close relationship between a TPD and matched model, and a TPD as a "fully" interacted TDH, we should be wary that a TDH is susceptible to the same issues we see in the matched model.

Time-Product Dummy

The Time-Product Dummy variable assigns a "dummy" or "indicator" variable to each "product" or model where each product is identified by a model number or a particular set of features.

Following the representation in de Haan et al. (2021), the time-product-dummy equation is:

$$\ln p_i^t = \alpha + \sum_{t=1}^{T} \delta^t D^t + \sum_{t=1}^{T} \gamma_i D_i$$

The TPD can accommodate additional fixed effects specific to a product. Any omitted variables from a hedonic specification, would contribute to a product specific fixed effect. TPD could also accommodate differences in coefficient values (e.g., differences in how a given feature contributes to the price of a car versus a truck). The TPD is the equivalent of a flexible, data-driven hedonic—Krsinich (2016) describes the TPD as a fully interacted TDH. However, it fails in the case of hedonic adjustment's reason for being, product entry and exit. The time-product dummy approach approximates a matched model index and only trivially includes "unmatched" observations. The TPD produces a similar index to the geometric matched model (Aizcorbe, 2014).

For a "good" hedonic specification, we would expect that the fixed effects for a product *i* would be approximately equal to the coefficient effects from the hedonic model. As de Haan points, the TPD is a special case of the TDH where:

$$\gamma_i = \sum_{k=1}^{K} \beta_k z_{ik}$$

Given the close relationship between a TPD and matched model, and a TPD as a "fully" interacted TDH, we should be wary that a TDH is susceptible to the same issues we see in the matched model.

Multilaterals with Hedonic Imputation

Ivancic, Diewert, and Fox (2011) introduced the GEKS formula and, more generally, sparked interest in multilateral approaches to address chain drift in price indexes. Chain drift can broadly be defined as divergence between the chained and fixed-base versions of a price index. The literature on chain drift focuses on the "stock" economic explanation for chain drift.[3] Here, consumers buy a product at price p_0 with a frequency of quantity q_0. The product goes on sale and quantity increases dramatically to q_1 and price decreases to p_1. Consumers stock up on the product the sale price, p_1, and satiate their demand over a longer time horizon than the measurement period. As such, even though the price returns to p0 in period 2 ($p_2=p_0$), quantity is significantly lower than the original amount demanded at the same price ($q_2<q_0$). Many price indexes, including the generally preferred superlative indexes, will show this as a permanent price decrease since the weight on the price increase is not symmetric with the weight on the price decrease. (Other indexes, such as the Jevons, an unweighted geometric index, would not show a permanent decrease). Other factors may lead to a divergence between chained and fixed-base index results especially when product turnover requires methods for product matching or grouping.

These methods have been combined with hedonic imputation in research beginning with de Haan and Krsinich (2014). De Haan and Daalmans (2019) discuss "single imputation," where only missing prices are imputed, and "double imputation" where a missing price and the observed price that corresponds to it in a price relative are imputed. They note that the double imputation may mitigate the effects of omitted variable bias.

Similarity Linking with Hedonic Imputation

Similarity linking has recently gained attention as a means of addressing chain drift in price indexes. Like multilateral methods, similarity linking first arose in the context of spatial price measurement but has

[3] See Diewert (2021).

been translated to intertemporal price indexes. As noted in *Modernizing the Consumer Price Index for the 21st Century* (National Academies, 2022), similarity linking has two advantages over multilateral indexes: first, the indexes satisfy the multiperiod identity test, and, second, are fully transitive, unlike rolling window extensions of multilateral indexes. The National Academies' report also suggests that hedonic imputation could be combined with similarity, and we explore that recommendation.

Similarity linking methods create chained price indexes where each period's price relative is a bilateral comparison between the given period and the prior period determined to be most "similar." The intuition being that price comparisons between periods with similar consumption patterns and weight distributions will reduce drift. The question arises of how to quantify "similarity."

Proposed methods include using the dissimilarity in predicted product shares between periods and the relative dispersion of the Laspeyres and Fisher indexes.

We introduce a method where the period specific hedonic regressions themselves are used to determine the similarity between periods using the test for regression model similarity proposed in Chow (1960). To determine the most similar link for a month t, we run Chow tests between t and each preceding period and take the period with the minimum Chow statistic as the link. Following the same process as other similarity linking procedures, index level for t, I^t, is calculated based on the bilateral price index, P, and most similar period to t, t^{min}:

$$I^t = I^{t^{min}} \times P(p^t, q^t, p^{t^{min}}, q^{t^{min}})$$

The Chow statistic measures how well a model estimated on a combination of two samples compares with models fitted on the samples individually. The Chow test consists of taking two sets of data—in our case, time period t and $t\text{-}a$—and estimating three regressions: one for each period and one where the data is combined into a single pool. The Chow statistic is then calculated based on the sum of squared errors from each regression, *SSE*, number of observations from each sample, *N*, and number of parameter estimates, *k*. These values produce the *F*-distributed Chow statistic:

$$F = \frac{(SSE_{Combo} - (SSE_t - SSE_{t-a}))/k}{(SSE_t - SSE_{t-a})/(N_t + N_{t-a} - 2k)}$$

We modify the typical Chow test to include a dummy variable for time in the combined regression, which, for t and $t\text{-}1$, would be equivalent to an adjacent period time dummy regression (this allows time periods to match based on similar coefficients even with aggregate price change). The time period with the lowest Chow statistic, t^{min}, is determined to be the most similar regression model to t and is selected as the link.

In addition to numerical advantages, similarity linking has significantly reduced computational requirements compared to multilateral indexes. Each new period of data is directly compared to each preceding period once and then an index is calculated. This means given w periods in an index window, w comparisons must be made to update a similarity index with $w\text{-}1$ similarity comparisons and one index calculation. This is a substantial reduction from GEKS-type indexes which require index comparisons on the order of w^2 for an w-length window.

6. Matched Model Indexes and Product Definition

The “matched model” index is the standard approach to measuring price change. Individual “models” are identified by either an indicator (for example, UPC or GTIN) or a set of specification values. The price for the same good is then compared from period-to-period. In a traditional, fixed sample, survey-based price index, product cycle effects are dealt with by comparing the price of a discontinued product with the price of a similar, successor product. When a replacement product is not considered comparable, the difference between the two items may be imputed. The item replacement and related-imputation process can have extremely large effects on a price index (See Williams, 2021).

In a scanner data context, the direct relationship between an exiting and entering good does exist as it does in a fixed sample survey. Products may be allowed to come and go from calculations as they enter and exit the market or remain on the market without any recorded sales. When “matched model” allows goods to fluidly enter and exit calculations, the “maximum overlap” approach to product turnover is used. However, this omits price change that may be introduced with model updates and allows for bias from product cycle effects. When working with scanner data, some researchers use the concept of a “product relaunch” to link old and new products together. Similarly, “product grouping” can be used so that multiple products can be grouped together and treated as one.

Here we investigate product definition in terms of aggregating transactions to a given model specification level. For example, following the Court specification, all observations that are all 180 inches long, 160 horsepower, and 2000 pounds would all be aggregated together to form a mean price and total quantity used in regression and matched model price index estimates. Once again, we use the specifications from Court, Ohta and Griliches, and our own specification. For a given specification, transactions are aggregated into an arithmetic mean price across all transactions meeting a given combination of variables and the total number of transactions as the quantity (with expenditure implied by the product of the mean price and total quantity). These indexes constitute a matched model index (without imputation) where a unique set of variable values for a given model specification constitutes a product definition.

7. Results

We reproduced the specifications used in Court (1939) and Ohta and Griliches (1976) for various forms of hedonic indexes namely pooled, adjacent period, and single period. Regression results for pooled version of the Court and Ohta and Griliches specifications are presented below in tables 2 and 3, respectively. The dummy variables for month have been excluded from both and nameplate has been excluded in table 3. The pooled result for our interacted specification is in the appendix. Results were similar to adjacent and single period coefficient estimates. Variables were generally significant and had the expected sign with the exception of “length” and “wheelbase” where these were negative (except for wheelbase in the interacted model).

Table 2: Regression results for the pooled Court model

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1.06E+01	1.57E-02	677.231	< 2e-16	***
Wheelbase	-2.84E-02	1.66E-04	-171.311	< 2e-16	***
Weight	6.31E-04	2.35E-06	268.856	< 2e-16	***
Horsepower	2.28E-03	7.41E-06	308.125	< 2e-16	***
MONTHS	-	-	-	-	-
	Multiple R-squared: 0.6887		Adjusted R-squared: 0.6884		

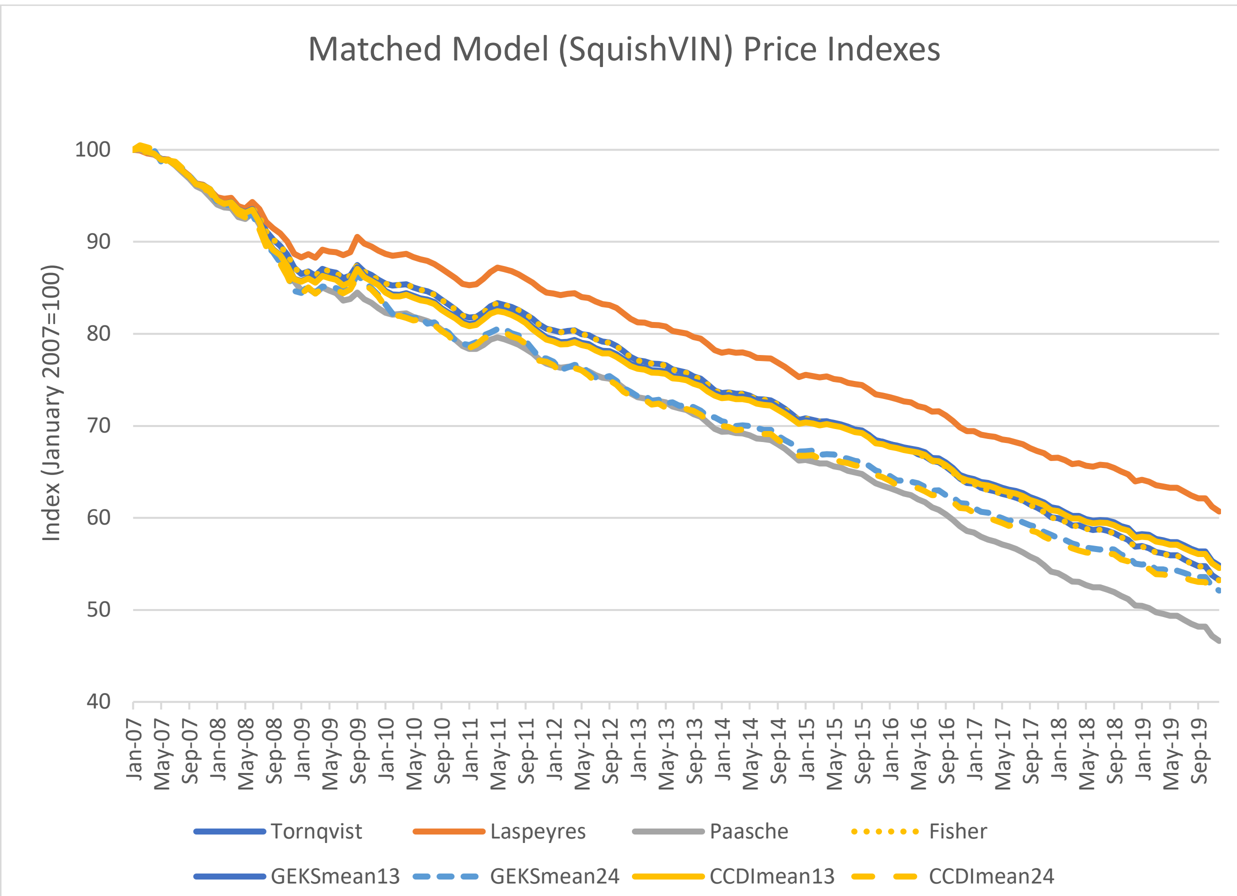
Table 3: Regression results for the pooled Ohta & Griliches model

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	9.75E+00	1.31E-02	742.181	< 2e-16	***
Length	-5.54E-03	6.38E-05	-86.912	< 2e-16	***
Weight	3.50E-04	1.98E-06	176.808	< 2e-16	***
Horsepower	1.12E-03	6.37E-06	176.371	< 2e-16	***
Cylinders 4	6.49E-02	8.02E-03	8.096	5.69E-16	***
Cylinders 5	-2.46E-02	8.63E-03	-2.854	0.004313	**
Cylinders 6	1.73E-01	8.20E-03	21.034	< 2e-16	***
Cylinders 8	4.29E-01	8.47E-03	50.681	< 2e-16	***
Cylinders 10	9.59E-01	1.05E-02	91.744	< 2e-16	***
Cylinders 12	7.56E-01	1.28E-02	58.97	< 2e-16	***
NAMEPLATE	-	-	-	-	-
MONTHS	-	-	-	-	-
	Multiple R-squared: 0.8717		Adjusted R-squared: 0.8715		

Matched Model Index Results

When using basic, matched model methods indexes drifted downward substantially. The results align with expectations from basic cost-of-living index theory: The Laspeyres and Paasche form upper and lower bounds (respectively), and the Törnqvist and Fisher are essentially equivalent. Applying multilateral methods does not address drift. This reinforces the finding from Williams and Sager (2019) that the index declines resulted from product cycles pricing patterns not weight-driven "chain drift." Interestingly, the multilateral indexes with longer window lengths showed more of a decline than shorter window lengths. The opposite of what we see in the hedonic imputation indexes. The final period chained bilateral Törnqvist and Fisher indexes are in between the 13- and 24-month multilateral indexes, suggesting little overall effect. Since no matches are being made across different versions of a product, longer windows for multilateral indexes will capture more sales for very old cars. This suggests extending the window will only worsen "drift." Extending the window length to 36 months resulted in bilateral comparisons with no matched observations.

The prevailing expectation may be that hedonic estimates would show lower indexes than conventional matched. To the contrary, we see that, of all the methodologies, matched model indexes produce the largest declines. This is consistent with similar research including Greenlees and McClelland (2010) and de Haan and Daalmans (2019). Matched model, maximum overlap price indexes show price change only for the same item so constant quality is maintained. These indexes also allow products to enter and exit calculations. They do not exhibit "quality bias" in the sense that price comparisons are made between goods of differing quality, which is often the motivation behind applying hedonic methods. However, the indexes are still subject to selection bias and product life cycle effects.

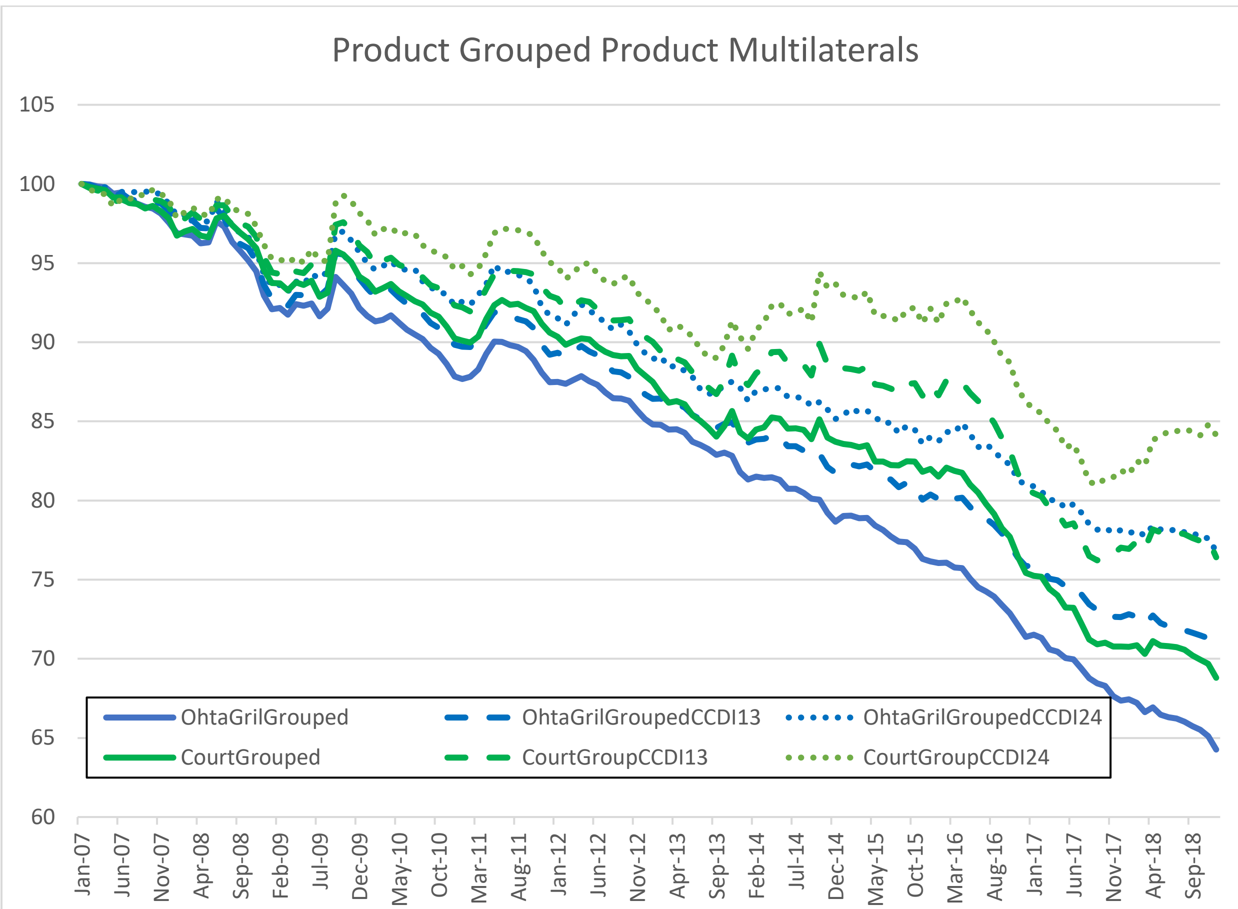


Product Grouping and Multilateral Indexes

As an alternative to hedonic imputation, cross-version price change can be measured by aggregating products with the same set of specification values and treating them as one product. As new iterations are introduced. Using the Court and Ohta and Griliches specifications to group products leads to indexes that decline much less than the matched model index based on a product identifier. Moreover, the application of multilateral formulas reduces the declines further. These indexes still do not represent plausible estimates for price change. For the decline of one product to be offset, it must have another exact match in terms of the specified features and continue to sell in the market. If an exact specification match does not exist, product cycle effects will bias the index.

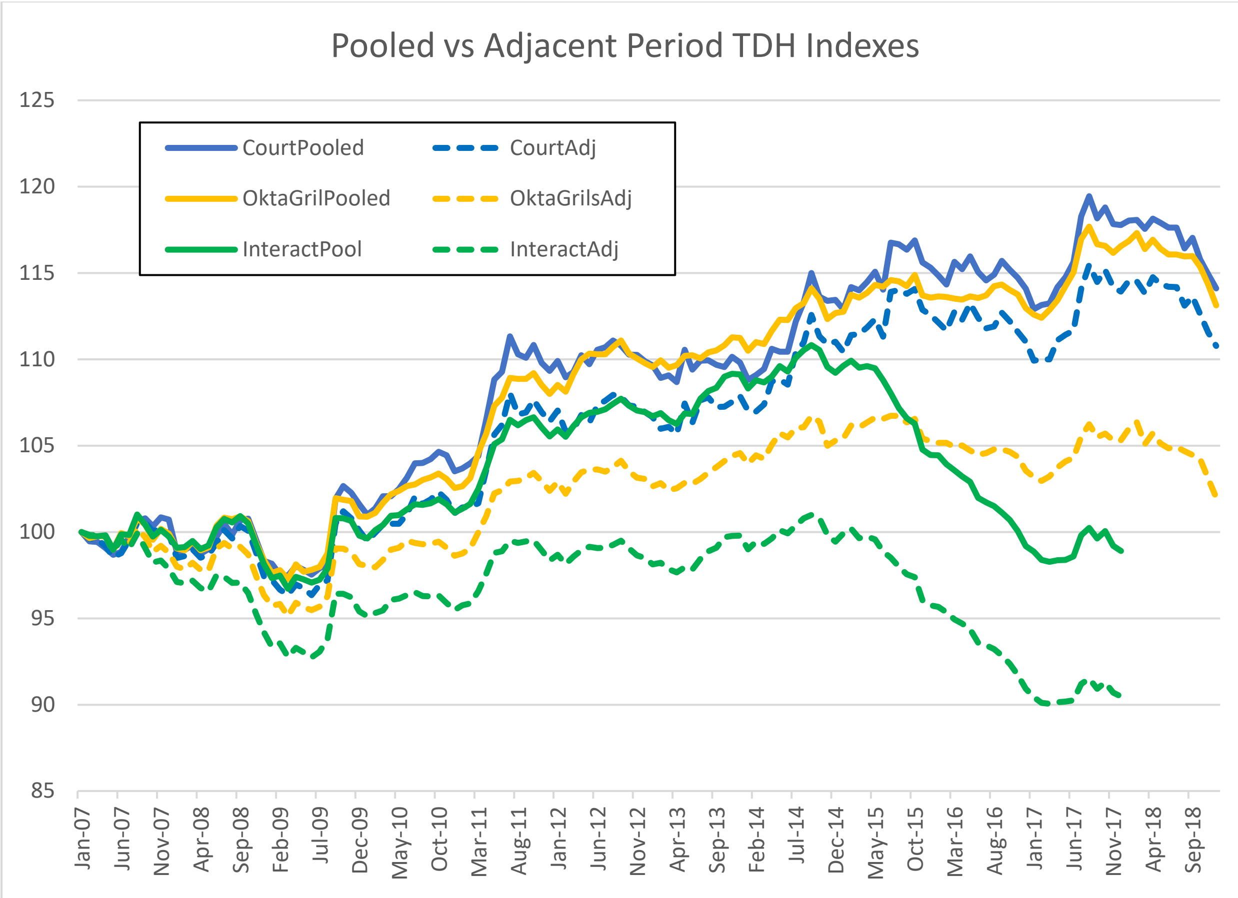
Product matching is often viewed as incidental to price index methods, however, our results show that making price comparisons across broader time horizons is essential for accurately measuring long-run price change. In other words, accumulations of short-term, same version price change do not result in accurate price measures—even when multilateral and similarity linking methods are applied. Hedonic and product grouping and matching methods are needed.

It is important to consider that many of the issues related to chain drift may arise as secondary effects that result from the method of product matching grouping or hedonic estimation applied to the data rather than a feature of the data in terms of a matched model.



Pooled and Adjacent Period Time Dummy Hedonic

Pooled regressions TDH were consistently higher than corresponding adjacent period indexes. Pooled regressions constrain coefficients to the same value over the entire period. The effect constrains the valuation of different features to remain the same over the entire period, which does not accommodate changes in consumer tastes. Pooled TDH also are also subject to revision as previous period values are reestimated with each additional month of data leading to revision.

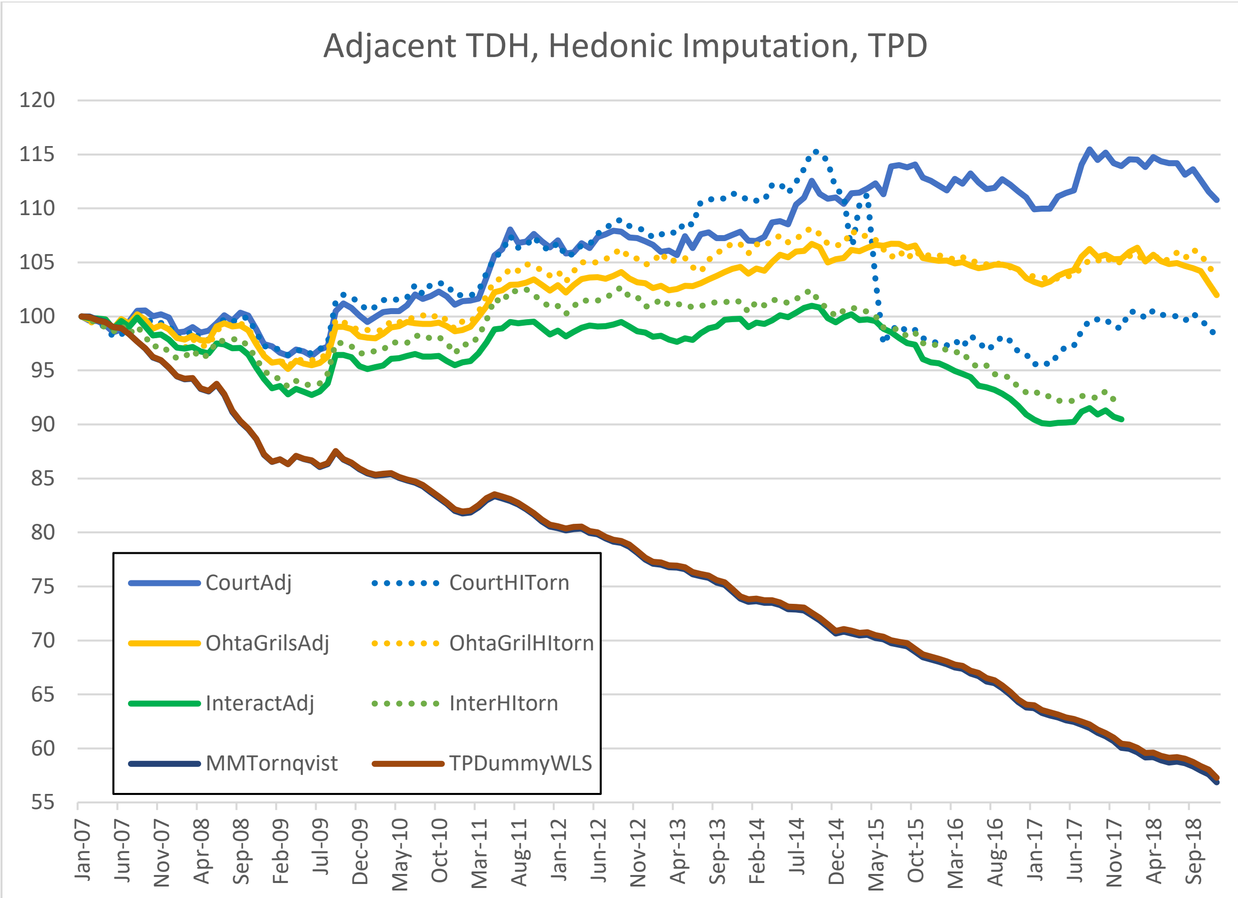


Bilateral Hedonic Imputation and Time-Product Dummy

Comparing these same adjacent period TDH indexes with their bilateral hedonic imputation counterparts shows little difference between the methods with an exception of period of divergence in the Court models. Both the Ohta and Griliches and interacted models were within a few percentage points of each other. The Court specification with hedonic imputation showed volatile behavior in 2015 that caused a divergence from its adjacent period counterpart. Shortly after the Court hedonic imputation index appears to stabilize and run close to parallel with the adjacent period index.

While the adjacent period and hedonic imputations appear plausible, there are still concerns that they may reflect product cycle bias.

Our results confirm the expectation that TPD and a geometric matched model index would perform similarly as the resulting indexes are extremely close.



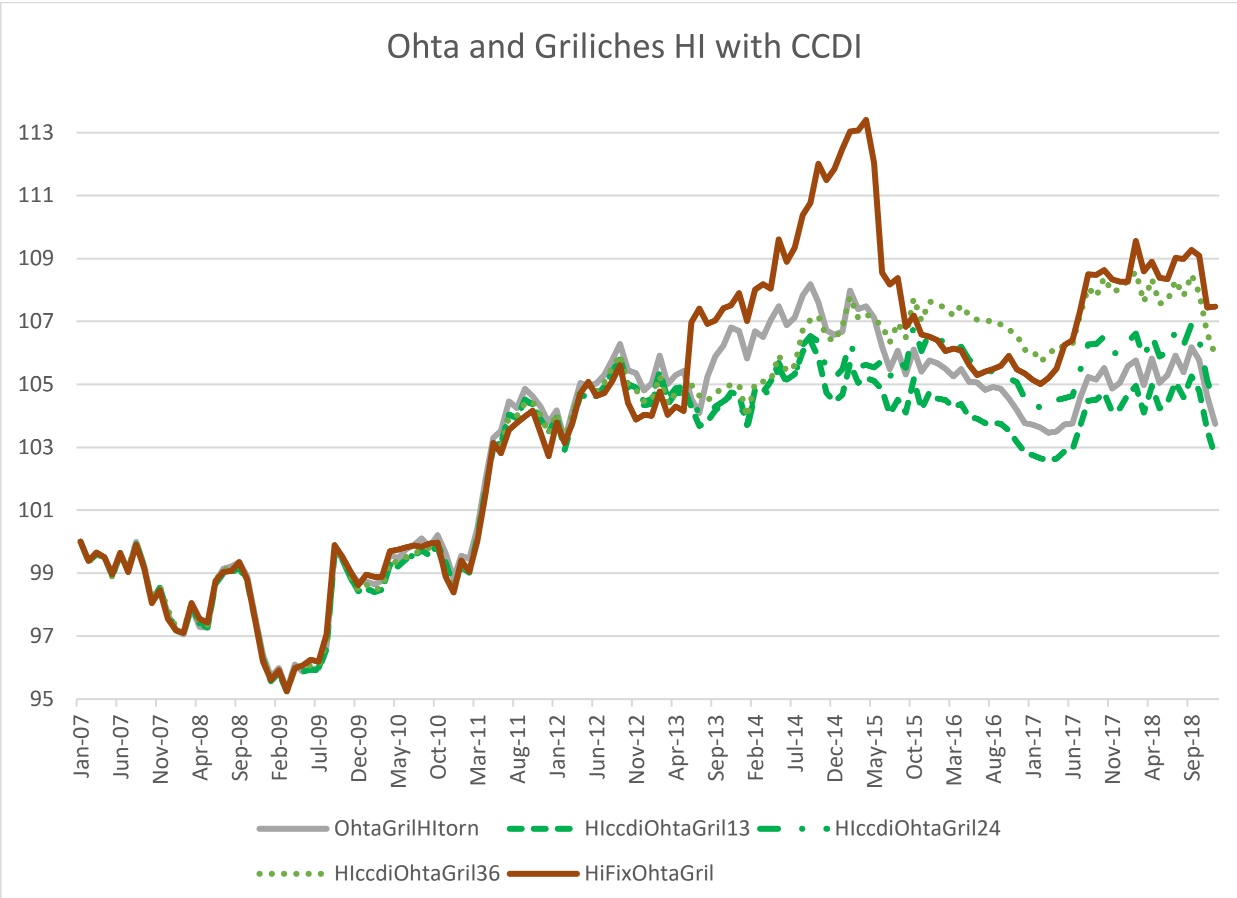
Hedonic Imputation with Multilateral Methods

The 13-month extension window had a downward effect compared to a bilateral hedonic imputation index (constructed on single period index imputation). The shorter window would reduce the occurrence of longer-run relatives compared to indexes with longer extensions, but it is unclear why it would lower an index below the bilateral hedonic imputation index.

Longer window multilaterals decline less than those with shorter windows. In the matched model indexes above this relationship was inverted with the 24-month window multilateral falling more than the 13-month. This suggests that the positive effects of extending the window are not related to addressing weight fluctuations that lead to drift, but, rather, increasing the representation of weight placed on longer-term, hedonically imputed price change. A fixed base, hedonic imputation index should not be sensitive to drift or product cycle effects, but the index will lose representivity over time as the base period set of products becomes less relevant. We construct a fixed base, Törnqvist index hedonic imputation which, over a 12-year span, is about 1.5% higher than the hedonic imputation CCDI (Caves-Christensen-Diewert-Inklaar index, a GEKS-type multilateral index based on Törnqvist bilateral comparisons) with a 36-month window.

To avoid product cycle effects in cars, an index must reflect price change across different iterations of goods (model years). A fully transitive index is not dependent on intervening periods, so within model year price change would not alter the long-run measurement of the index. However, full period multilaterals are difficult to calculate because of product turnover and computational demand.

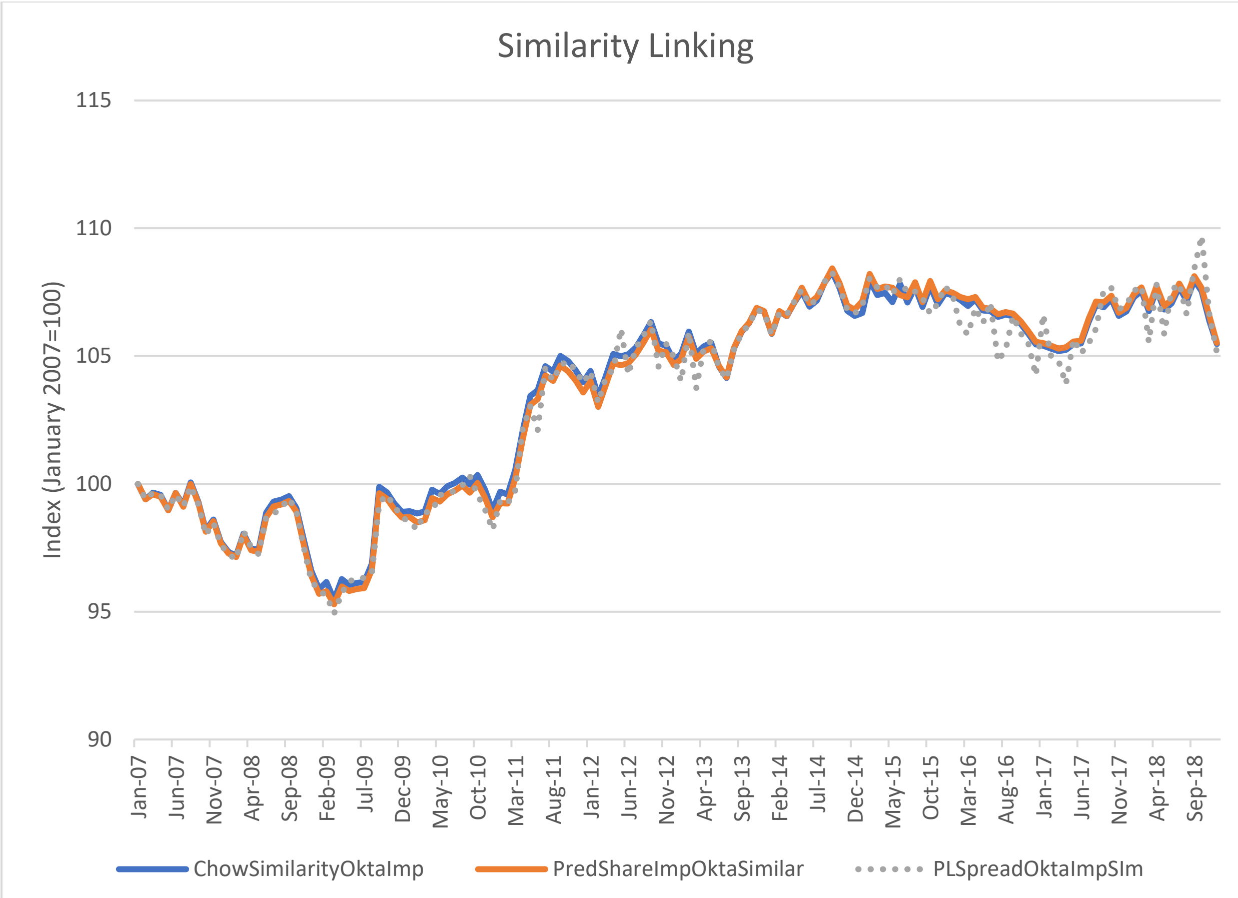
Moreover, they lead to revisions of prior months which are not acceptable for the publication of many official statistics. Extension methods can lead to indexes that are nearly transitive, but longer windows are preferred to better capture long-run price comparisons.



Similarity Linking with Hedonic Imputation

The three methods of similarity linking combined with hedonic imputation all produced similar results. The Chow similarity and predicted share methods were highly correlated. The similarity link indexes without hedonic imputation (matched model indexes with similarity linking) showed large declines. The case mirrors the results of applying multilateral methods to matched model indexes: Without product matching or hedonic imputation to offset product cycle effects and capture price change with model updates, indexes will decline to implausible levels.

Our indexes using similarity linking showed index results comparable to a CCDI index with a 36-month extension window. However, the most similar month was typically the proceeding month with 105 of the 143 periods tested selecting the month prior as the most similar. Unlike GEKS-type multilateral, similarity linking does not necessarily force a comparison over a longer-time horizon. If changes are incremental or if the most "similar" link remains the previous month even after a pricing regime changes, similarity linking may not address aspects of the product cycle.



8. Conclusion

Hedonic estimates have often been used to impute the prices for entering and exiting products. Hedonic estimates may also be used to estimate long-run price relatives, which allow better measurement of price change across product cycles. Product cycle effects have generally been neglected and the focus has been on "quality bias." In matched model indexes, quality bias emerges as a secondary effect from the use of product matching as a means of addressing product cycle issues including price change with model updates and price discrimination.

Previous research has also found that multilateral indexes with hedonic imputation tend to fall less when a longer extension window is used. We find evidence that this is mostly due to the additional influence long-term, cross-product cycle relatives have in multilaterals with longer extended windows.

Estimates from hedonic imputation can be used with similarity linking methods. Like other multilateral methods, similarity linking without product replacement or hedonic imputation does not remedy product cycle effects. Using regression model similarity as a method for linking produces similar results to other multilateral methods but with greater simplicity and less computational demand.

Works Cited

Aizcorbe, Ana M. 2014. *A Practical Guide to Price Index and Hedonic Techniques.* Oxford: Oxford University Press.

Aizcorbe, Ana, Benjamin Bridgman, and Jeremy Nalewaik. 2010. "Heterogeneous car buyers: A stylized fact." *Economic Letters* 109 (1): 50-53. doi:10.1016/j.econlet.2010.08.003.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841-890. doi:10.2307/2171802.

Brown, Craig, and Anya Stockburger. 2006. "Item Replacement and Quality Change in Apparel Price Indexes." *Monthly Labor Review* 129 (12): 35-45. https://www.bls.gov/opub/mlr/2006/12/art3full.pdf.

Chow, Gregory C. 1960. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." *Econometrica* 28 (3): 591-605. doi:10.2307/1910133.

Colwell, Peter F., and Gene Dilmore. 1999. "Who Was First? An Examination of an Early Hedonic Study." *Land Economics* 75 (4): 620-626. doi:10.2307/3147070.

Court, A.T. 1939. "Hedonic Price Indexes with Automotive Examples." *General Motors Corporation, Automobile Demand.*

Cowling, Keith, and John Cubbin. 1972. "Hedonic Price Indexes for United Kingdom Cars." *Economic Journal* 82 (327): 963-978.

de Haan, Jan, and Jacco Daalmans. 2019. "Scanner Data in the CPI: the imputation CCDI Index revisited." *16th meeting of the Ottawa Group.*

de Haan, Jan, Rens Hendriks, and Michael Scholz. 2016. "A Comparison of Weighted Time-Product Dummy and Time Dummy Hedonic Indexes." Graz Economics Papers.

Diewert, Erwin. 2018. "“Scanner Data, Elementary Price Indexes and the Chain Drift Problem”." (Vancouver School of Economics) Discussion Paper 20-07.

Diewert, W. Erwin. 2019. "Quality Adjustment and Hedonics: A Unified Approach." Vancouver School of Economics.

Feenstra, Robert C. 1995. "Exact Hedonic Price Indexes." *The Review of Economics and Statistics* 77 (4): 634-653.

Goodman, Allen C. 1998. "Andrew Court and the Invention of Hedonic Price Analysis." *Journal of Urban Economics* 44 (2): 291-298. doi:10.1006/juec.1997.2071.

Greenlees, J., and R. McClelland. 2010. "Superlative and Regression-Based Consumer Price Indexes for Apparel Using U.S. Scanner Data." St. Gallen, Switzerland: Conference of the International Association for Research in Income and Wealth.

Greenlees, John S., and Robert McClelland. 2011. "Does Quality Adjustment Matter for Technologically Stable Products? An Application to the CPI for Food." *The American Economic Review* 101 (3): 200-205. https://www.jstor.org/stable/29783739.

Griliches, Zvi. 1961. "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change." In *The Price Staistics of the Federal Government*, 173-96. NBER.

Ivancic, Lorraine, W. Erwin Diewert, and Kevin J. Fox. 2011. "Scanner data, time aggregation and the construction of price indexes." *Journal of Econometrics* 161 (1): 24-35. doi:10.1016/j.jeconom.2010.09.003.

Johnson, David S, Stephen B. Reed, and Kenneth J Stewart. 2006. "Price measurement in the United States: a decade after the Boskin Report." *Monthly Labor Review* 129 (5): 10-19. https://www.bls.gov/opub/mlr/2006/05/art2full.pdf.

Konny, Crystal G., Brendan K. Williams, and David M. Friedman. 2019. "Big Data in the U.S. Consumer Price Index: Experiences and Plans." In *Big Data for Twenty-First Century Economic Statistics*, edited by Katharine G. Abraham, Ron S. Jarmin, Brian Moyer and Matthew D. Shapiro. University of Chicago Press. http://www.nber.org/chapters/c14280.

Krsinich, Frances. 2014. "The FEWS index: Fixed effects with a window splice." *Meeting of the Group of Experts on Consumer Price Indices.* Geneva, Switzerland. https://unece.org/sites/default/files/datastore/fileadmin/DAM/stats/documents/ece/ces/ge.22/2014/New_Zealand_-_FEWS.pdf.

National Academies of Sciences, Engineering, and Medicine. 2022. *Modernizing the Consumer Price Index for the 21st Century.* Washington, DC: The National Academies Press.

Ohta, Makoto, and Zvi Griliches. 1976. "Automobile Prices Revistited: Extensions of the Hedonic Hypothesis." In *Household Production and Consumption*, edited by Nestor E. Teleckyj, 325-398. NBER.

Pakes, Ariel. 2003. "A Reconsideration of Hedonic Price Indexes with an Application to PC's." *The American Economic Review* 93 (5): 1578-1596. https://www.jstor.org/stable/3132143.

Reinsdorf, Marshall B., Paul Liegey, and Kenneth Stewart. 1996. "New Ways of Handling Quality Change in the U.S. Consumer Price Index." *BLS Working Paper 276.* https://www.bls.gov/osmr/research-papers/1996/pdf/ec960040.pdf.

Silver, Mick, and Saeed Heravi. 2007. "The Difference between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes." *Journal of Business & Economic Statistics* 25 (2): 239-246. http://www.jstor.org/stable/27638928.

Triplett, Jack E. 1969. "Automobiles and Hedonic Quality Measurement." *Journal of Political Economy* 77 (3): 408-417. doi:10.1086/259524.

Triplett, Jack. 2006. *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes.* Paris: OECD Publishing.

Williams, Brendan. 2021. "Twenty-One Years of Adjustments for Quality Change in the US Consumer Price Index." *Group of Experts on Consumer Price Indices.* Online: UNECE. https://unece.org/sites/default/files/2021-05/Session_3_US-BLS_Paper_0.docx.

Williams, Brendan, and Erick Sager. 2019. "A New Vehicles Transaction Price Index: Offsetting the Effects of Price Discrimination and Product Cycle Bias with a Year-Over-Year Index." *BLS Working Papers* (Working Paper 514). https://www.bls.gov/osmr/pdf/ec190040.pdf.

Appendix

Table 4: Pooled Interacted Regression

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	8.5590	0.0268	319.715	< 2e-16	***
BASE..ins..	0.0058	0.0002	33.998	< 2e-16	***
Length..ins..	-0.0002	0.0001	-2.464	0.013759	*
weight	0.0003	0.0000	45.374	< 2e-16	***
horsepower	0.0016	0.0001	28.577	< 2e-16	***
AWDdummy	0.0806	0.0013	62.736	< 2e-16	***
displacement	-0.2948	0.0044	-66.646	< 2e-16	***
height	-0.0155	0.0002	-67.134	< 2e-16	***
MPGCity	0.0216	0.0011	20.094	< 2e-16	***
MPGHwy	-0.0168	0.0011	-14.834	< 2e-16	***
HybrDummy	0.4583	0.0179	25.607	< 2e-16	***
torque	0.0011	0.0000	121.504	< 2e-16	***
I(horsepower/weight)	18.0400	0.1710	105.457	< 2e-16	***
Make1	-0.2496	0.0027	-91.931	< 2e-16	***
Make2	-0.1644	0.0470	-3.497	0.000471	***
Make3	-0.1594	0.0038	-41.437	< 2e-16	***
Make4	-0.1441	0.0027	-52.95	< 2e-16	***
Make5	-0.1405	0.0021	-67.696	< 2e-16	***
Make6	-0.1396	0.0055	-25.335	< 2e-16	***
Make7	-0.1328	0.0022	-59.761	< 2e-16	***
Make8	-0.1301	0.0039	-33.104	< 2e-16	***
Make9	-0.1283	0.0025	-50.53	< 2e-16	***
Make10	-0.1239	0.0022	-55.747	< 2e-16	***
Make11	-0.1148	0.0066	-17.365	< 2e-16	***
Make12	-0.0870	0.0032	-26.788	< 2e-16	***
Make13	-0.0521	0.0038	-13.8	< 2e-16	***
Make14	-0.0333	0.0021	-15.912	< 2e-16	***
Make15	-0.0215	0.0025	-8.516	< 2e-16	***
Make16	-0.0211	0.0021	-9.951	< 2e-16	***
Make17	-0.0194	0.0024	-8.062	7.56E-16	***
Make18	-0.0066	0.0029	-2.305	0.021175	*
Make19	0.0632	0.0022	28.216	< 2e-16	***
Make20	0.1381	0.0059	23.425	< 2e-16	***
Make21	0.1722	0.0030	57.081	< 2e-16	***
Make22	0.1824	0.0039	47.164	< 2e-16	***

Make23	0.1907	0.0029	65.505	< 2e-16	***
Make24	0.2104	0.0124	16.971	< 2e-16	***
Make25	0.2244	0.0032	69.286	< 2e-16	***
Make26	0.2343	0.0112	20.872	< 2e-16	***
Make27	0.2479	0.0026	94.48	< 2e-16	***
Make28	0.2561	0.0043	60.248	< 2e-16	***
Make29	0.3073	0.0026	117.379	< 2e-16	***
Make30	0.3261	0.0030	108.382	< 2e-16	***
Make31	0.3711	0.0026	144.678	< 2e-16	***
Make32	0.4131	0.0033	125.356	< 2e-16	***
Make33	0.4226	0.0028	153.396	< 2e-16	***
Make34	0.5864	0.0278	21.129	< 2e-16	***
Make35	0.8271	0.0180	45.862	< 2e-16	***
Make36	0.8463	0.0030	279.909	< 2e-16	***
cylinders3	-0.0625	0.0084	-7.429	1.10E-13	***
cylinders4	-0.1437	0.0041	-35.445	< 2e-16	***
cylinders5	-0.2308	0.0042	-54.936	< 2e-16	***
cylinders6	-0.1201	0.0027	-44.113	< 2e-16	***
cylinders10	0.2364	0.0056	42.428	< 2e-16	***
cylinders12	0.2428	0.0080	30.451	< 2e-16	***
BODYSTYLEconvertible	0.1467	0.0017	85.684	< 2e-16	***
BODYSTYLEcoupe	0.0261	0.0014	19.003	< 2e-16	***
BODYSTYLEhatchback	0.0209	0.0015	14.213	< 2e-16	***
BODYSTYLEwagon	0.0823	0.0017	48.44	< 2e-16	***
MPGCity:HybrDummy	0.0018	0.0004	4.759	1.95E-06	***
MPGHwy:HybrDummy	-0.0102	0.0005	-19.042	< 2e-16	***
HybrDummy:torque	0.0005	0.0000	15.657	< 2e-16	***
displacement:MPGHwy	0.0207	0.0003	65.344	< 2e-16	***
displacement:MPGCity	-0.0160	0.0004	-40.57	< 2e-16	***
horsepower:MPGCity	0.0001	0.0000	14.697	< 2e-16	***
horsepower:MPGHwy	-0.0002	0.0000	-68.501	< 2e-16	***
weight:MPGHwy	0.00001	0.0000	13.83	< 2e-16	***
weight:MPGCity	0.0000	0.0000	-0.219	0.827012	