

# Cataloging Growth: A Re-Evaluation of 1900–1990\*

Verónica Bäcker-Peral Benjamin Wittenbrink  
MIT MIT

October 25, 2026

## Abstract

Measuring real GDP growth requires distinguishing changes in prices from changes in product quality and composition, yet systematic quality adjustment of price indexes is unavailable for much of the twentieth century. We construct a new quality-adjusted price index for U.S. housefurnishings and apparel goods using 5.1 million product listings from Sears catalogs, 1900–1990. We use large language models to extract product information and estimate hedonic price schedules with detailed product characteristics, allowing us to infer annual changes in the cost of living. The resulting cost-of-living index implies substantially lower goods inflation than conventional deflators, and consequently implies much faster real economic growth: between 1900 and 1990, real consumption of housefurnishings and apparel grew by a factor of 102 using our index, compared with a factor of 24 using standard goods deflators. The gap between our index and canonical ones is largest before World War II and suggests that productivity growth peaked earlier in the twentieth century than conventional estimates indicate.

*JEL Classification:* E01, N11, N12, O47, C43.

---

\*We would like to thank Abhijit Banerjee and Dave Donaldson for invaluable guidance and mentorship. We are very grateful to Daron Acemoglu, David Atkin, Ricardo Caballero, Amy Finkelstein, Joe Hazell, Jack Kelly, Russell Legate-Yang, Atif Mian, Emi Nakamura, Christina Romer, David Romer, David Weinstein, Iván Werning, Chris Wolf and workshop participants at MIT for comments and suggestions. We thank the George and Obie Shultz Fund at MIT for financial support.

# 1 Introduction

The period of most transformative growth in modern US history is also amongst the least understood. Economic historians have emphasized the century from roughly 1870 to 1970 as an exceptional episode in American economic growth (Gordon, 1999; Abramovitz and David, 2000; Gordon, 2016). There is less agreement about how or when that growth took place. Different studies place the peak of U.S. productivity growth either before or after World War II (Field, 2003; Gordon, 2013; Bakker et al., 2019; Crafts and Woltjer, 2021; Field, 2023), while qualitative accounts emphasize substantial technological progress during both periods (Gordon, 2016). Progress on this question has been slow in large part because high-quality price data are scarce for most of this period. As Gordon (1999) puts it, “the history of price changes from 1914 to 1947 [is] the black hole where little is known.” This paper sheds light on that black hole.

Economists have long recognized that accurate price measurement is essential for welfare comparisons. A 1915 Bureau of Labor Statistics (BLS) report lamented that “men did not attempt to measure changes in the level of prices until after they had learned to measure such subtle magnitudes as the weight of the atmosphere, the velocity of sound, fluctuations of temperature, and the precession of the equinoxes” (United States Bureau of Labor Statistics, 1915). Rapid inflation during World War I made price measurement especially salient, and the BLS responded by introducing the Consumer Price Index (CPI), which aims to be a nationally representative measure of the cost of living. How prices are measured shapes much of what we know about growth. Nominal U.S. Gross Domestic Product grew by a factor of nearly 500 over the twentieth century, but the CPI attributes 44% of average annual nominal growth to changes in the cost of living alone.<sup>1</sup>

There is a long-standing concern, however, that the CPI has historically overstated inflation by failing to fully account for changes in the quality of consumer goods and services (Boskin et al., 1998). The quality of goods available to consumers has improved dramatically in the last century. A refrigerator in 1900 was a large wooden box, much of it taken up by a compartment that held blocks of ice to keep the rest cold. By the 1930s, households could purchase electric refrigerators, which continued to improve over the century as they added features such as ice makers and water dispensers. Mattresses, often filled with cotton felt, sea moss, or horsehair in the early 1900s, improved markedly over the century through better spring systems and the introduction of foam and latex for greater support and comfort. Innovations in textiles led to softer fabrics, machine-washable clothing, and elastic materials. More broadly, almost every aspect of the American consumer landscape was transformed (Gordon, 2016). In the 1990s, the BLS revised the CPI methodology to account for quality change in a subset of goods and extrapolated some of these revisions back to 1978, but for much of the earlier period, systematic quality adjustment was more limited.<sup>2</sup>

This paper constructs a new quality-adjusted price index for U.S. housefurnishings<sup>3</sup> and apparel over the twentieth century that is consistent with modern BLS methods. We digitize a nearly complete run of Sears general merchandise catalogs from 1900 to 1990, containing approximately 5.1 million product listings. For each product category and year, we estimate hedonic price schedules from detailed product characteristics, then aggregate these estimates into a quality-adjusted price index. We present two main results: first, the overall rise in the standard of living over the twentieth century was much larger than conventional estimates suggest. Using existing price indexes for housefurnishings and apparel, real consumption in these sectors grew by a factor of 24 between 1900 and 1990; using our quality-adjusted index, it grew by a factor of 102. We argue that this revision reflects two limitations of the BLS’s

<sup>1</sup> We obtain data on long-run GDP from the Macrohistory Database (Jordà et al., 2017). We obtain historical CPI data from the Minneapolis Federal Reserve Bank.

<sup>2</sup> A notable exception is new automobile consumption, for which the BLS began explicit quality adjustment starting in 1966.

<sup>3</sup> Our definition of housefurnishings includes recreational and entertainment goods, like televisions and radios, which was the BLS convention for most of the twentieth century.

historical price-index methodology. First, for much of the twentieth century, the absence of explicit product identifiers made it difficult for BLS field workers to track identical products over time. Historical reports suggest that field workers often substituted higher-quality alternatives for older products without adequately adjusting for differences in quality, thereby overstating price increases. Second, as extensively documented in the literature, the BLS’s historical focus on matched-model methods introduced additional bias by restricting price comparisons to products that survived across periods. In contrast, modern hedonic methods allow prices to be estimated in subsequent periods for all base-period products, reducing this source of selection bias (Griliches, 1961; Boskin et al., 1998; Gordon, 1990; Pakes, 2003). Carried through to real growth, the difference implies that housefurnishings and apparel consumption rose about 1.6 percentage points faster per year than existing deflators record.

Our second main result is that a substantial share of 20th century growth in housefurnishings and apparel took place before World War II. Average annual growth in these sectors was 3.3% from 1900 to 1941 and 3.6% from 1946 to 1990. Replacing the conventional goods deflator with our quality-adjusted Sears index reverses this pattern. We estimate average annual real consumption growth in housefurnishings and apparel of 6.1% for 1900-1941 and 4.8% for 1946-1990.

The Sears catalogs provide an unusually rich source for historical price measurement. Sears was the largest retailer in the US for most of the 20th century and published two general merchandise catalogs per year, in the spring and fall, as well as a Christmas catalog. These catalogs listed nationally posted prices for a broad range of consumer goods and described products in considerable detail. We digitize a nearly complete collection of these catalogs from 1900 to 1990 and use large language models to extract structured product records from the catalog pages.

Building on a long tradition of empirical work, we use hedonic methods to construct a cost-of-living index for US housefurnishings and apparel. Hedonic methods are especially useful in settings like ours for which we have rich data on product prices and characteristics but very limited data on quantities of goods sold. For 168 categories of goods, we use an LLM to identify detailed product characteristics that are featured in the Sears catalogs. We then use an LLM to extract information on each of these characteristics for all of the products in our sample. Using this detailed product characteristic data, we estimate hedonic regressions for each of the product categories in our sample for each five-year rolling window. Like the BLS, we track identical products across periods whenever possible and use hedonic methods to adjust for quality differences when products must be replaced. We aggregate these estimates into a price index for housefurnishings and apparel using nationally representative weights. Under general conditions outlined in Pakes (2003), the resulting quality-adjusted index approximates the true change in the cost of living for these sectors to first order. If anything, this interpretation understates the welfare gains from quality improvement, since the second-order bias goes in the direction of overstating the cost of living.

We also assess whether Sears is informative about broader movements in retail prices for consumer goods. Sears was the largest retailer in the US for most of the twentieth century, accounting for about 6% of retail sales, and it sold a broad range of consumer durables and household goods, serving consumers nationwide. Its share of retail output in the post-World War II period was comparable to that of Amazon today. As a robustness check, we construct a parallel quality-adjusted price index using J.C. Penney catalogs for 1969–1982. J.C. Penney was also a major national retailer, but served a higher-end segment of the market. The two catalog-based indexes are similar over their period of overlap and yield similar hedonic coefficients. Moreover, when we replicate the BLS methodology for the period of 1978-1990,<sup>4</sup> we obtain very similar inflation estimates. This comparison suggests that the main results are not specific to Sears and are informative about broader movements in retail prices for consumer goods.

Our hedonic approach allows us to explore the observable dimensions along which quality changed over the twentieth century. We find that the observable quality of housefurnishings and apparel goods in our sample improved by about 1.9% annually over 1900-1990. A significant share of this growth

<sup>4</sup> For this period, we assume that the BLS was effectively able to track identical or highly similar products over time.

was driven by the introduction of new features, technologies and materials. About 70% of annual quality improvement over 1905-1990 was driven by new features that were not present in goods listed in 1900-1905. Electrification was a major margin along which quality changed: by 1939, more than half of the refrigerators and washing machines in the Sears catalog were electric, and after the war the share quickly reached 100%. Although quality changes are typically positive, economic theory predicts that during severe recessions, quality should fall as households substitute away from luxury goods (Fajgelbaum et al., 2011; Jaimovich et al., 2019). This is exactly what we find. At the onset of the Great Depression, measured quality declined for three years, as Sears shifted toward smaller and lower-quality versions of products it had previously offered. This episode provides further evidence that the quality upgrading we measure using hedonic regressions is capturing real shifts in the economy.

Our Sears-based estimates cover housefurnishings and apparel but not other goods, housing, or services. To assess the aggregate implications, we combine our index with existing price indexes for these other sectors and compare aggregate real consumption growth under two alternative housefurnishings and apparel deflators: the existing BLS index and our quality-adjusted Sears index.<sup>5</sup> Holding other sectors fixed, our adjustment is still consequential for changes in total consumption. Using the standard BLS index, U.S. real total consumption grew by a factor of 15.6 between 1900 and 1990; using ours it grew by a factor of 20.4 — an average difference of 0.3 percentage points per year. Assuming that annual price index bias for total consumption expenditures is equal to that of the private domestic economy and applying our results to the total factor productivity (TFP) estimates in Bakker et al. (2019), we find that the period of highest TFP growth was earlier than the authors’ estimates suggest.

**Related Literature** This paper contributes to the literature on measurement of long-run growth in the U.S. and other advanced economies. Gordon characterizes the twentieth-century transformation of the US economy as “one big wave” that peaked in the mid century (Gordon, 1999, 2013). Subsequent work has debated whether productivity growth was highest before, during, or after World War II (Field, 2003; Gordon, 2013; Bakker et al., 2019; Crafts and Woltjer, 2021; Field, 2023). Our results offer a new perspective on this debate. After accounting for quality improvements in apparel and housefurnishings, we find that TFP growth was highest in 1929–1941, although estimated growth in 1948–1960 and 1960–1973 was only slightly lower. More broadly, all three periods exhibit higher productivity growth than preceding and subsequent decades. Our adjustment thus shifts the estimated peak earlier relative to Bakker et al. (2019), who identify 1960–1973 as the period of fastest productivity growth. Our estimates also revise upward TFP growth in the 1910s and 1920s. Taken together, these findings are consistent with Gordon’s characterization of “one big wave” of productivity growth centered around World War II, which had largely subsided by the mid-1970s.

Catalog data has long been recognized as a consistent source for historical price measurement. Rees and Jacobs (1961) used Sears and Montgomery Ward catalogs to construct price indexes for clothing and home furnishings for 1890–1914, following the conventional BLS methodology of tracking repeated products and substituting the most similar alternative when a product disappears. This approach, still standard today, is well established to be biased relative to the true change in the cost of living (Pakes, 2003; Bils, 2009; Nakamura and Steinsson, 2012; Ehrlich et al., 2026). In an influential study, Gordon (1990) hand-collected data on several categories of durable goods, including appliances from Sears catalogs, for 1947–1983 and estimated hedonic regressions to construct a quality-adjusted index, finding that cumulative real growth in Producer’s Durable Equipment was about 3 times higher than the official index implies. Relative to Gordon, we expand the scope of time and products. Our implied revision is more modest: for 1947–1983, we find that cumulative real consumption growth for housefurnishings and apparel was 1.4 times higher than the official index implies. The difference reflects that some consumer goods saw less quality improvement over this period than investment durables like machinery

<sup>5</sup> We do not include investment goods, which Gordon (1990) shows saw much more dramatic quality improvement than retail goods, so price index bias for total GDP is likely higher than for total consumption.

and vehicles, so the official consumer-goods index carries less quality bias.

The hedonic approach treats goods as bundles of characteristics and estimates how market prices vary with them. Early work adjusted automobile prices for quality change (Court, 1939; Griliches, 1961, 1971), Rosen (1974) provided the canonical equilibrium interpretation, and subsequent work clarified when hedonic coefficients can be used to construct price indexes (Triplett, 1986, 2006; Pakes, 2003). Nordhaus (1996) studied the quality-adjusted price of lighting over a very long horizon, showing that conventional price indexes can miss large gains from technological change. Our contribution is to extend the hedonic tradition to a far broader historical setting, using millions of Sears product descriptions to estimate quality-adjusted indexes across many categories of consumer goods over nearly the entire twentieth century.

The main alternatives to the hedonic approach rely on quantity or expenditure data. One strand imposes structure on preferences to construct exact price indexes (Diewert, 1976; Sato, 1976; Vartia, 1976), extended by Feenstra (1994) to allow entry and exit and widely applied to measure the gains from new varieties (Hausman, 1997; Broda and Weinstein, 2006; Redding and Weinstein, 2020) and the growth that imputation misses when entrants displace incumbents (Aghion et al., 2019). The other uses Engel curves: Hamilton (2001) and Costa (2001) infer CPI bias from shifts in food expenditure shares, while Bils and Klenow (2001) recover unmeasured quality growth in durables from the tendency of richer households to buy higher-quality goods. Both require expenditure shares or quantities, which catalog data lacks, so neither is feasible in our setting.

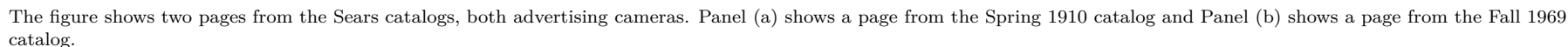
**Outline** Section 2 describes how we construct our primary dataset of Sears catalogs. Section 3 presents a simple framework for constructing a cost-of-living index using hedonic methods and develops our empirical methodology. Section 5 presents our estimated price index for consumer goods and several robustness checks. Section 6 discusses the implications of our findings for measuring real GDP growth and productivity, and Section 7 concludes.

## 2 Data

Our primary data consist of 180 Sears general merchandise catalogs, published bi-annually, spanning 1900 to 1990. We also include the Sears Christmas books for 1945 to 1990, as well as the Sears Home Annual catalogs for 1986 to 1990. For these years, many appliances were included only in the Home Annual catalog, and so it is important to add these to our dataset.<sup>6</sup> We augment the Sears series with J.C. Penney catalogs (1969–1982) to broaden coverage and to assess robustness to retailer composition.

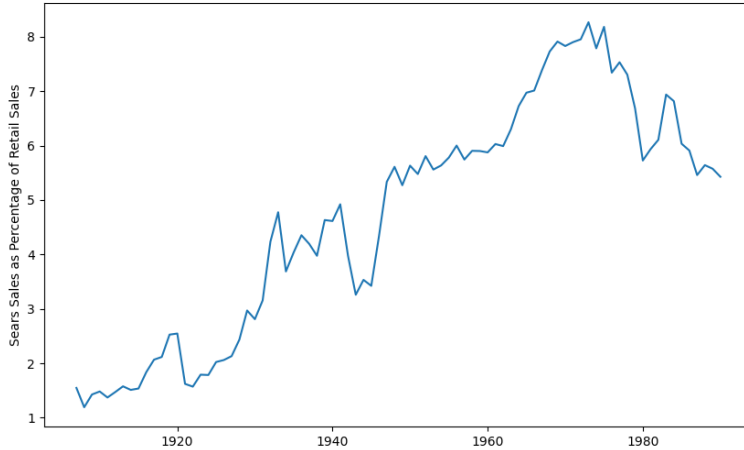
<sup>6</sup> From the semiannual catalogs, we are missing the 1906 Fall and 1907 Spring catalogs, which we could not find in archival collections.

(b) 1969



We process catalog data at the page level and extract structured product listings using a large language model (LLM) with a standardized schema, as in [Bäcker-Peral et al. \(2025\)](#); [Asirvatham et al. \(2026\)](#). [Figure 1](#) shows an example of a page from the Sears catalog in 1910 and another from 1969, both featuring cameras. In both years, the listings are detailed, providing prices, model-level descriptions, and information on size, accessories, and other features of each distinct camera. To extract information from catalog images, we use Gemini 3 Pro, which is a state-of-the-art LLM that is highly effective at processing text from images. Each catalog contains roughly 1,000-1,500 pages, yielding 5,121,232 Sears product records.<sup>7</sup> For each listing we record: posted price, product title, textual description, shipping weight (when reported), brand (when reported), and a product image when present. We also ask the LLM to classify each product into a 1992 6-digit Harmonized System (HS) code, which is the classification used by the World Customs Organization.<sup>8</sup> The text field is particularly valuable: it contains detailed claims about materials, performance, dimensions, included accessories, and warranties, which are often precisely the characteristics that evolve as quality changes. In a second round of processing, we classify each product into a United Nations Standard Products and Services Code (UNSPSC) product category using OpenAI’s GPT 5-mini. Relative to the HS classification, UNSPSC groups products by their purpose rather than material and is more detailed in its classification schema of consumer goods. In the remainder of the paper, if not otherwise specified, when we classify products into “categories” we use their UNSPSC classification. Finally, we ask the LLM to classify products which are accessories to the primary item being listed. For example, Sears often sold items like batteries or replacement parts on the same page as the item they were paired with. For our main analysis, we exclude these accessories. In [Section A.2](#), we provide additional details on the digitization methodology.

**Figure 2:** Sears Sales as a Percent of Retail Sector Output



The figure is a line plot of total Sears retail sales as a percent of U.S. retail sector output. For the post-1947 period, we use retail sector gross output data from the BEA. Before this, we use retail sector value added from historical Census data, and re-normalize it to match the BEA level for 1947–1957. This effectively assumes that the input-output share of retail was equal to the 1947–1957 level for the first half of the twentieth century. We exclude sales from Sears subsidiaries like Allstate Insurance.

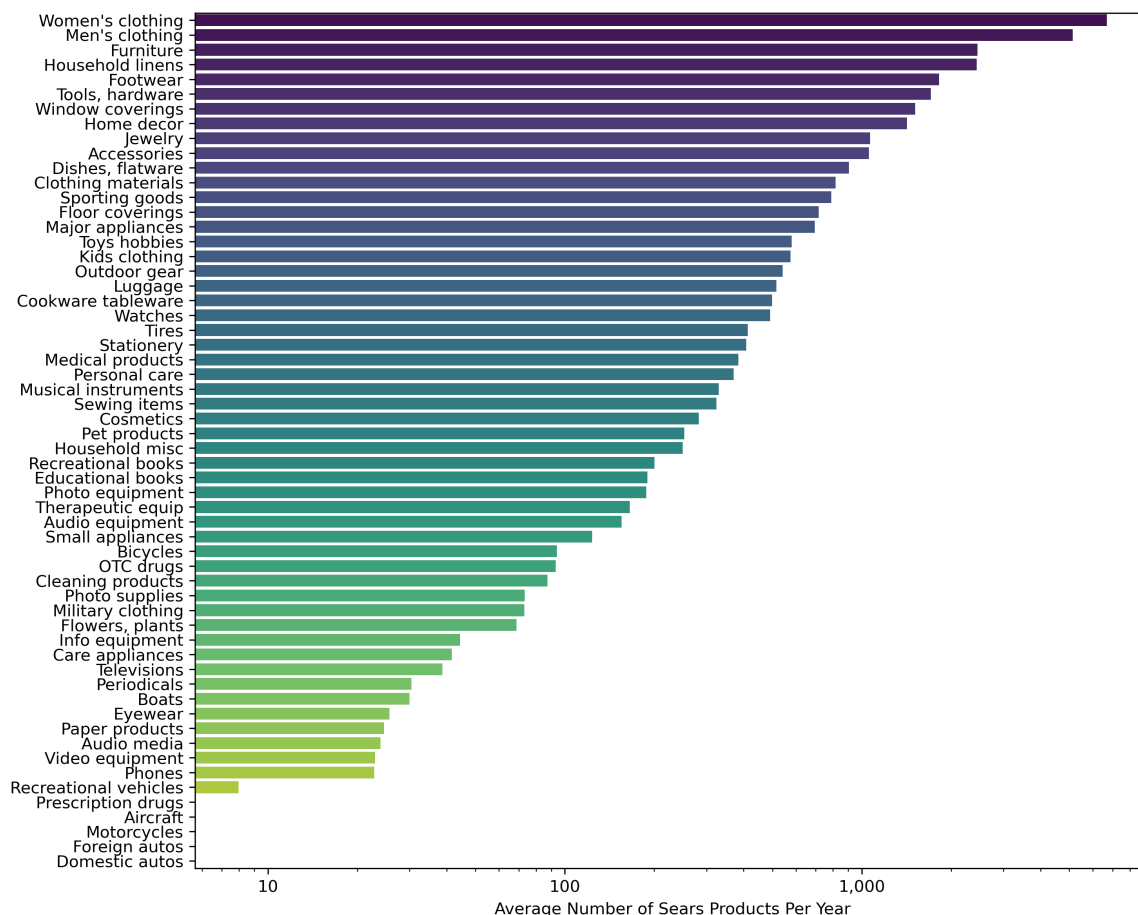
<sup>7</sup> For reference, the BLS collects about 80,000 quotes for goods and services per month. About 23% of these are for goods, excluding food and energy. Therefore, if the BLS had collected at this rate for the period of 1900-1990, we expect that the BLS collected about 20 million prices — about four times larger than our dataset. In practice, the BLS hand-collected a much smaller number of quotes in the pre-World War II period, so it is likely that our dataset is of approximately the same size or larger for these periods.

<sup>8</sup> There are 5,022 HS codes for 1992 recognized by the World Customs Organization. In our sample, we observe products spanning 3,430 HS codes.

## 2.1 Historical importance of Sears

Sears was a major retail institution for much of the twentieth century, but it began as a mail-order business aimed at customers far from large urban retail centers. Richard W. Sears started by selling watches by mail in the mid-1880s and, after partnering with Alvah C. Roebuck and relocating the business to Chicago, expanded into a general merchandise catalog that offered an unusually broad variety of standardized, mass-produced goods at posted prices. In an era when many households—especially in rural areas—faced limited local selection, mail-order catalogs functioned as a central purchasing channel: consumers could shop from home and have goods delivered through the rapidly expanding national logistics system.<sup>9</sup> By the late 1970s, Sears reportedly printed on the order of hundreds of millions of catalog copies annually and served tens of millions of customers, with catalog sales in the billions of dollars. **Figure 2** plots Sears retail sales as a share of U.S. retail sector gross output over our sample period. Sears’ footprint rises from roughly 1.5 percent of retail sales in the early twentieth century to more than 6 percent for several decades following World War II, before declining in the 1980s. For reference, **Amazon, Inc. accounted for 6.6% of US retail sales in 2024**, so Sears can be thought of as “the Amazon of its day.”

**Figure 3: Sears Coverage of Consumer Goods**



The figure shows a bar plot indicating the average number of products listed annually by Sears within each of the Personal Consumption Expenditure (PCE) categories for consumer goods. The x-axis uses a logarithmic scale.

<sup>9</sup> Two infrastructure developments were particularly important for making this model scalable: the maturation of the railroad network (which lowered the cost of long-distance freight) and the introduction of Rural Free Delivery in the late 1890s (which made home delivery feasible for rural addresses).

The Sears catalog data do not span the full CPI consumption basket. In particular, Sears excludes housing, services, food, fuel and vehicle consumption, all of which are central CPI components.<sup>10</sup> For that reason, we interpret our results as evidence on inflation and quality change for *housefurnishings and apparel*, rather than as a direct measure of the aggregate cost of living.

Within the consumer goods sectors that we focus on, however, Sears offered a large selection of goods of different varieties. **Figure 3** shows the wide coverage of products in the Sears catalogs. We map products from the Sears catalogs to the 58 consumer goods categories from the BEA Personal Consumption Expenditures (PCE) series and plot the average number of products that appear each year in the Sears catalog for each PCE category. All PCE categories are covered, with the exception of aircraft, motorcycles, automobiles, and prescription drugs. The figure indicates that the Sears catalog featured an especially large collection of apparel, furniture and other housefurnishings.

### 3 Empirical framework

We are interested in how household welfare changed over the twentieth century. Let household  $h$ 's utility in period  $t$  be

$$u_t^h = U^h(\mathbf{q}_t^h), \quad (3.1)$$

where  $\mathbf{q}_t^h$  denotes its vector of consumption. Rather than compare utility levels directly, we can express welfare in expenditure units using the expenditure function

$$e^h(\mathbf{p}, u) \equiv \min_{\mathbf{q}} \left\{ \mathbf{p} \cdot \mathbf{q} : U^h(\mathbf{q}) \geq u \right\}. \quad (3.2)$$

Let  $x_t^h \equiv e^h(\mathbf{p}_t, u_t^h)$  denote household  $h$ 's expenditures in period  $t$ .

We assume that preferences are homothetic, so that utility can be represented by a function homogeneous of degree one. Under this normalization, the expenditure function is linear in utility:

$$e^h(\mathbf{p}, u) = u e^h(\mathbf{p}, 1). \quad (3.3)$$

Consequently, changes in utility can be expressed exactly as changes in the expenditure required to attain utility, holding prices fixed. In particular,

$$\begin{aligned} \frac{u_t^h}{u_{t-1}^h} &= \frac{e^h(\mathbf{p}_t, u_t^h)}{e^h(\mathbf{p}_t, u_{t-1}^h)} \\ &= \frac{x_t^h}{x_{t-1}^h} \cdot \frac{e^h(\mathbf{p}_{t-1}, u_{t-1}^h)}{e^h(\mathbf{p}_t, u_{t-1}^h)}, \end{aligned} \quad (3.4)$$

where the second line uses the definition of household expenditure.

Define the household-level cost-of-living index as

$$COLI_t^h \equiv \frac{e^h(\mathbf{p}_t, u_{t-1}^h)}{e^h(\mathbf{p}_{t-1}, u_{t-1}^h)}. \quad (3.5)$$

This is a cost-of-living index in the sense of **Konüs (1939)**: it measures the change in the minimum expenditure required to maintain household  $h$  at its base-period utility.

<sup>10</sup> Sears did, however, sell **mail-order houses** from 1908-1942.

We can aggregate across households using base-period expenditure weights,  $\lambda_h \equiv x_{t-1}^h / \sum_g x_{t-1}^g$ . Then, to a first-order approximation,

$$\begin{aligned} \sum_h \lambda_h \frac{u_t^h}{u_{t-1}^h} &= \sum_h \lambda_h \frac{x_t^h / x_{t-1}^h}{COLI_t^h} \\ &\approx \frac{\sum_h \lambda_h (x_t^h / x_{t-1}^h)}{\sum_h \lambda_h COLI_t^h} \\ &= \frac{Y_t / Y_{t-1}}{COLI_t}, \end{aligned} \quad (3.6)$$

where aggregate nominal consumption expenditure is  $Y_t = \sum_h x_t^h$ , and the aggregate cost-of-living index is defined as

$$COLI_t \equiv \sum_h \lambda_h COLI_t^h. \quad (3.7)$$

In practice, the cost-of-living index cannot be directly observed because it depends on household utility functions, which are unknown. A standard approximation evaluates the cost of the base-period consumption bundle at current-period prices. Let  $\mathbf{q}_{t-1}^h$  denote household  $h$ 's base-period consumption bundle and  $\mathbf{p}_t$  the vector of current-period prices. The corresponding Laspeyres index is

$$L_t \equiv \sum_h \lambda_h \frac{\mathbf{q}_{t-1}^h \cdot \mathbf{p}_t}{\mathbf{q}_{t-1}^h \cdot \mathbf{p}_{t-1}}. \quad (3.8)$$

By revealed preference,  $COLI_t \leq L_t$ , because the Laspeyres index holds the base-period bundle fixed and therefore ignores the possibility that households substitute away from goods whose relative prices rise. Moreover, for small price changes, the Laspeyres index provides a first-order approximation to the true cost-of-living index. By Shephard's lemma, the derivative of the expenditure function with respect to prices equals the vector of Hicksian quantities. A first-order expansion of the expenditure function therefore implies that the expenditure required to maintain base-period utility is approximately equal to the cost of purchasing the base-period bundle at current prices. We provide the formal derivation in [Section A.3](#).

The Laspeyres index can equivalently be written as an expenditure-weighted average of product-level price relatives:

$$L_t = \sum_{i \in \Omega_{t-1}} s_{i,t-1} \frac{p_{it}}{p_{i,t-1}}, \quad (3.9)$$

where  $\Omega_{t-1}$  denotes the set of base-period products and

$$s_{i,t-1} \equiv \frac{\sum_h p_{i,t-1} q_{i,t-1}^h}{\sum_h x_{t-1}^h}$$

is the aggregate base-period expenditure share on product  $i$ .

The empirical difficulty is that [Equation \(3.9\)](#) requires prices for the same goods in adjacent periods. This requirement is difficult to satisfy in settings with substantial product turnover. Products enter and exit the Sears catalog frequently, and continuing product categories often change in quality, style, materials, and other observable characteristics. As a result, the period- $t$  price of a base-period product is often unobserved.

### 3.1 Hedonic quality adjustment

A natural solution to product turnover is to compare each product with its closest substitute in the following period. However, if goods improve over time, directly comparing the prices of successive

products would conflate quality improvements with inflation, leading us to overstate increases in the cost of living. A standard way to address this problem is through hedonic methods, which use the relationship between prices and observable product characteristics to estimate the value of differences in quality (Griliches, 1961; Pakes, 2003; Ehrlich et al., 2026).

Formally, suppose that product  $i$  belongs to a product category  $j(i) \in \mathcal{J}$  and is described by a vector of characteristics  $X_{it} \in \mathcal{X}^{j(i)}$ . In a competitive equilibrium, Rosen (1974) shows that prices will be a function of product characteristics,

$$p_{it} = P_{j(i),t}(X_{it}).$$

For example, for a refrigerator, the price may be a function of its size, color, and the various features it contains, like an ice machine or a water dispenser. We construct a vector of characteristics,  $X_{it}$ , for every product, with the use of an LLM. First, we identify approximately one hundred category-specific attributes for 168 categories of consumer goods. These include product sizes, materials, technological features and aesthetic attributes. Then, for every item belonging to each of these 168 categories, we extract the value of each of these attributes, whenever it can be determined from the product description. We provide further details for how we extract structured attributes in Section A.5.

Assuming that log prices are linear in product characteristics, we can estimate the value of each characteristic using the following hedonic regression,

$$\log p_{it} = \delta_t^{j,\tau} + X'_{it}\beta^{j,\tau} + \varepsilon_{ijt}, \quad t \in [\tau - 2, \tau + 2]. \quad (3.10)$$

We estimate Equation (3.10) separately for every product category  $j$  and year  $\tau$ , where  $\delta_t^{j,\tau}$  are year and season fixed effects. In our baseline specification, we pool the estimation of Equation (3.10) over a five-year window,  $[\tau - 2, \tau + 2]$ , in order to capture stable hedonic coefficients rather than fleeting trends. However, our results are robust to using different sized windows. The hedonic coefficients,  $\beta^{j,\tau}$ , capture the implicit prices of observable product characteristics. The fit of the hedonic regression is high, with a median out-of-sample  $R^2$  of 0.73, which is comparable to the  $R^2$  obtain with hedonic methods using modern data (Ehrlich et al., 2026). The distribution of  $R^2$  across product categories and years is shown in Figure A.1.<sup>11</sup> Under the assumptions in Section A.3, these coefficients provide first-order approximations to the welfare-relevant value of changes in product characteristics (Griliches, 1961; Pakes, 2003; Ehrlich et al., 2026).

A commonly observed product attribute over time is shipping weight. We choose to exclude shipping weight from our baseline hedonic regression because it is confounded with quality — in any given year, the heaviest products also tend to be the most expensive, but over time, firms aim to reduce the weight of goods without reducing their quality. As a result, our hedonic regression will capture a positive coefficient on shipping weight and will record declines in weight as declines in quality, when in reality, these may be driven by increases in quality from the introduction of lighter materials. Instead, we include product dimensions in the hedonic regression, which are more likely to be qualities that consumers value.

One practical obstacle is that the matrix  $X_{it}$  contains approximately one hundred attributes for every product, and these attributes may be strongly correlated. We therefore estimate Equation (3.10) via a ridge regression where the year and season fixed effects are left unpenalized. However, our results are robust to instead selecting attributes with a Least Absolute Shrinkage and Selection Operator (LASSO) regression and estimating Equation (3.10) via OLS.

Given the hedonic coefficients from Equation (3.10), we can infer next-period prices for every product by adjusting the observed price of its closest substitute for differences in characteristics. In particular, suppose we observe a product  $i$  of type  $j$  at time  $t - 1$ . We identify the most similar product of

<sup>11</sup> Another test of model fit is to predict log price *changes* as a function of hedonic characteristics. For this exercise, we obtain a median out-of-sample  $R^2$  of 0.40, which is again consistent with the modern literature (Ehrlich et al., 2026). The distribution of  $R^2$  across product categories and years with log price change as the dependent variable is shown in Figure A.2

type  $j$  observed at time  $t$  based on the text descriptions of the products. To do so, we map product descriptions into embedding space using OpenAI’s `text-embedding-3-small` model, which maps text into a high-dimensional vector space designed to capture semantic similarity. We measure similarity using cosine similarity, the normalized inner product of two vectors, with higher values indicating more similar descriptions.

Let  $k(i)$  be the most similar product of type  $j$  observed at time  $t$ . We estimate that  $i$ ’s price in the subsequent period is given by,

$$\hat{p}_{i,t} = p_{k(i),t} \exp \left( (X_{i,t-1} - X_{k(i),t})' \hat{\beta}^{j,t} \right). \quad (3.11)$$

Intuitively, we take the observed price of the closest substitute and adjust it by the estimated value of the differences in product characteristics. Whenever  $i$  and its closest substitute are observationally identical in characteristic space,  $X_{i,t-1} = X_{k(i),t}$ , we take the substitute’s price directly. When the closest substitute differs along observable dimensions, however, we adjust the price accordingly. This allows us to separate changes in prices from changes in observable quality, even when the exact same product cannot be identified in adjacent periods.

We then construct a quality-adjusted Laspeyres index,

$$\widehat{COLI}_t = \sum_{i \in \Omega_{t-1}} s_{i,t-1} \cdot \frac{\hat{p}_{i,t}}{p_{i,t-1}} \quad (3.12)$$

where  $\Omega_{t-1}$  denotes the set of base-period products and  $s_{i,t-1}$  denotes nationally representative expenditure weights.<sup>12</sup> This way, we are able to estimate the contribution to the change in the cost-of-living for every product in the base period, even if it no longer exists a year later.

To reiterate, the index in Equation (3.12) is a first order approximation of the object of interest. With richer quantity data, one could construct closer approximations to the cost of living. For example, if product-level prices and quantities were observed in adjacent periods, one could use Törnqvist-style indexes that average expenditure shares across periods. Such indexes allow expenditure shares to respond to relative price changes and therefore provide a better approximation to substitution than a base-weighted Laspeyres index (Ehrlich et al., 2026).

Our setting does not permit that approach. Sears catalogs provide rich information on product descriptions and posted prices, but not transaction quantities. The same limitation prevents us from implementing exact demand-based price indexes of the kind used in Broda and Weinstein (2006) and Redding and Weinstein (2020), which require stronger information on quantities and expenditure shares, and allow researchers to estimate gains from access to new product varieties. Our empirical strategy instead uses the comparative advantage of the Sears data: detailed product descriptions over a long historical period. Hedonic regressions allow us to decompose these long-run price changes into a quality change component and a cost-of-living change component. To the degree that households benefited from expansion in varieties over this period, in addition to changes in quality, our approach will therefore underestimate the rise in welfare.

### 3.2 Expenditure weights

The index in Equation (3.12) weights each product-level price change by its base-period expenditure share,  $s_{i,t-1}$ . In practice, these shares are unobserved in our data. We observe catalog prices and characteristics, but, with rare exceptions, we observe neither the quantities Sears sold nor—what the index ultimately requires—national expenditure at the level of individual products.

Our strategy is to recover  $s_{i,t-1}$  in two steps, relying on nationally representative survey data wherever it exists and on the catalog-based weights only where it does not. National consumer surveys

<sup>12</sup>In practice, we aggregate goods using a geometric average *within* product categories, and a weighted arithmetic average *across* product categories, consistent with the BLS methodology (U.S. Bureau of Labor Statistics, 2018).

report representative expenditure weights, but only at a moderate level of aggregation. The Bureau of Labor Statistics (BLS) organizes spending into three major retail categories, indexed by  $k$ —household goods, apparel, and entertainment goods—and roughly 60 subcategories, indexed by  $\ell$ . We map each subcategory  $\ell$  to one or more UNSPSC product codes and assign every Sears listing to a subcategory through its code. We take the expenditure weight on each subcategory,  $w_{\ell,t-1}^{\text{BLS}}$ , directly from the national source. These weights are nationally representative by construction and require no assumption about Sears.

We construct expenditure weights from historical consumer expenditure surveys and published BLS CPI weights. For 1900–1924, we use expenditure shares from the *Cost of Living in the United States* survey, conducted in 1917–1919. Following historical BLS methodology, for 1925–1929 we adjust the broad expenditure shares for apparel and housefurnishings to the average of the 1917–1919 and 1935–1936 surveys, and for 1930–1934 to the 1935–1936 shares, while retaining the earlier detailed category weights within each group (U.S. Bureau of Labor Statistics, 1941). For 1935–1946, we use detailed expenditure shares from the *Study of Consumer Purchases in the United States*, conducted in 1935–1936. From 1947 onward, we use historical CPI expenditure weights published by the BLS, updating them as new expenditure benchmarks become available.<sup>13</sup>

National sources do not resolve spending below the subcategory level, so we cannot read product-level shares from any representative survey. We instead use the catalog itself, under the assumption that within a subcategory, expenditure is allocated proportionally across all catalog products in that subcategory. Formally,

$$s_{i,t-1} = \frac{1}{N_{\ell(i),t-1}} \times w_{\ell(i),t-1}^{\text{BLS}} \quad (3.13)$$

where  $\ell(i)$  denotes the subcategory to which product  $i$  is mapped,  $N_{\ell(i),t-1}$  is the number of catalog products in subcategory  $\ell$  at  $t-1$ . We take geometric averages of price changes within BLS product categories, and arithmetic averages of price changes across BLS, consistent with modern BLS aggregation practices.

This construction rests on the assumption that within a subcategory, the Sears product mix is representative of what households actually purchased. This assumption is not directly testable at the product level, but we can provide suggestive evidence in its favor.

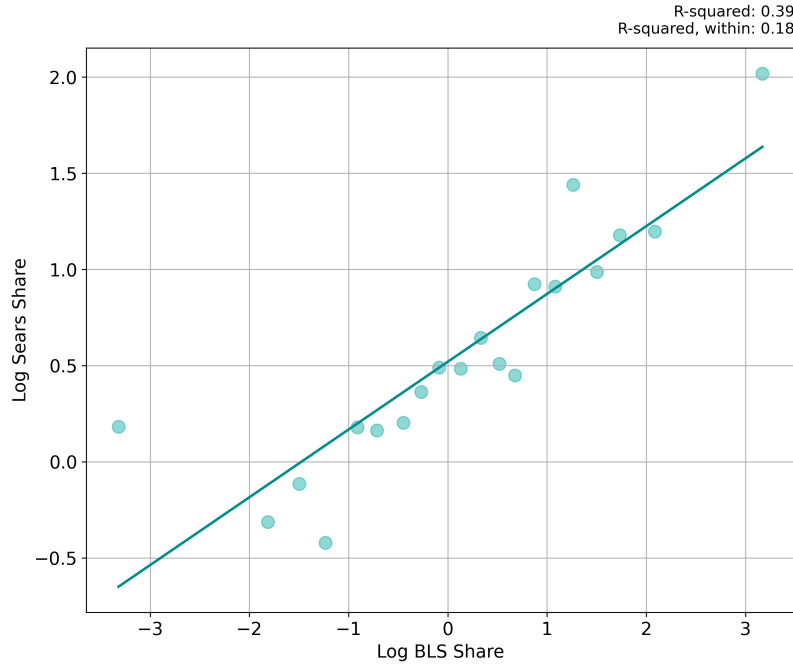
**Evidence on representativeness.** To assess the assumption that *within* BLS categories Sears is representative of national consumption, we check whether it aligns with the BLS expenditure weights *across* categories. We define the Sears share of subcategory  $\ell$  within major category  $k$  as its share of distinct listings,

$$\text{Sears Share}_{\ell,k,t} = \frac{\text{Num Products}_{\ell,k,t}}{\sum_{\ell'} \text{Num Products}_{\ell',k,t}}. \quad (3.14)$$

Analogously, we define  $\text{Sears Share}_{k,t}$  for each major category. At the level of major categories, the catalog is *not* representative: [Figure A.3](#) shows that Sears over-weights apparel and household goods relative to the BLS. This is exactly why we anchor cross-category weights to the BLS rather than to the catalog. Within major categories, however, the catalog tracks national spending well. [Figure 4](#) plots  $\log \text{Sears Share}_{\ell,k,t}$  against  $\log \text{BLS Share}_{\ell,k,t}$ , controlling for category  $k$  and year fixed effects; the strong positive relationship indicates that, conditional on a major category, the internal composition of the catalog mirrors national expenditure patterns. This supports the assumption that within BLS subcategory, the products that are listed more frequently in Sears will also be those that had higher national expenditure shares.<sup>14</sup>

<sup>13</sup>For years without new expenditure data, we hold weights fixed at the relevant benchmark, consistent with historical BLS practice.

<sup>14</sup>Ideally, we would want the slope in [Figure 4](#) to be one, such that products with higher expenditures have proportional weights in the Sears catalog. Because the slope is less than one, these results suggest that we will under-weight more

**Figure 4:** Binscatter of Sears Product Shares vs. CPI Subcategory Weights

Binscatter with 100 bins. The vertical axis is  $\log \text{Sears Share}_{\ell,k,t}$ , the share of subcategory  $\ell$  within major category  $k$  in the Sears catalog in year  $t$ , defined in (3.14). The horizontal axis is the corresponding  $\log \text{BLS Share}_{\ell,k,t}$  reported by the BLS. Both axes residualize on category  $k$  and year  $t$  fixed effects. We restrict to years in which BLS surveys were fielded.

## 4 Comparison of Our Methodology and the BLS

The methodology described in Section 3.1 is intended to replicate the modern BLS methodology closely. However, earlier versions of the CPI followed a notably different methodology which evolved over the twentieth century. In this section, we describe the major changes to the CPI over our period of study and discuss how we implement both the historical and modern methodology in our data.

### 4.1 Overview of the BLS Methodology Over the 20th Century

In this section, we describe the major changes to the BLS methodology over the period we study, focusing on the housefurnishings and apparel sectors.

#### 4.1.1 1900–1913

Before 1913, the BLS did not publish an official cost-of-living index. However, the [Minneapolis Fed historical price index](#) series uses the price index constructed by [Rees and Jacobs \(1961\)](#).

For housefurnishings and apparel, [Rees and Jacobs \(1961\)](#) uses catalog data from Montgomery Ward and Sears, so the underlying data are very similar to our study. However, there are important methodological differences. [Rees and Jacobs \(1961\)](#) identified between one and two items in the 1918 Sears catalog for about half of the listed housefurnishings and apparel items in the *Cost of Living in the United States* survey (fielded 1917–1919). The two items were selected “so that the price of one was about equal to the average expenditure per article for all families covered in the 1918 budget study and

---

popular goods.

the price of the other, to the average expenditure per item for families in the income class whose average expenditure was lowest.” The authors also prioritized goods that were “durable and serviceable rather than fashionable.” After selecting about one hundred specific 1918 housefurnishings and apparel items, the authors followed each item backward annually, selecting in each year the closest substitute for the original good. When no comparable product could be found, the authors dropped the original one and replaced it with a similarly priced item that could be linked backward.

While a tremendous effort to construct price data for this period, [Rees and Jacobs \(1961\)](#) has several limitations relative to our study. First, they use only a small subset of the vast catalog data available over the period of study. Second, even if the items selected in 1918 were somewhat representative of contemporaneous consumer expenditures, it is unlikely that they would be representative of expenditures a decade or two earlier.

#### 4.1.2 1914–1941

Following rapid inflation during World War I, the BLS fielded the *Cost of Living in the United States* survey (1917–1919), which collected expenditure data from 12,000 families in 92 cities in 42 states. Regular collection of retail prices began in 1917, but the BLS extended the price index back to 1913 using a combination of historical retail and wholesale price indexes ([United States Bureau of Labor Statistics, 1966](#); [Reed, 2014](#)). The BLS fielded a second expenditure survey in 1934–1936, to address changing household consumption patterns, which collected expenditure information from approximately 14,500 families in 42 cities ([United States Bureau of Labor Statistics, 1966](#)).

The CPI tracked a fixed set of representative goods over time, collecting prices periodically from retailers. For housefurnishings and apparel, BLS agents visited stores in person twice a year, generally obtaining four price quotations per article in each city. Given that the BLS tracked about 80 housefurnishings and apparel items and 32 cities, we estimate that they obtained about 20.4 thousand price quotes each year; in Sears, we observe on average about 19.6 thousand housefurnishings and apparel items in each year for this period. When a product disappeared, agents were instructed to select a replacement that was as similar as possible in quality and grade ([U.S. Bureau of Labor Statistics, 1923](#)).

However, contemporaneous reports suggest that product specifications were relatively imprecise, and it was difficult to identify comparable goods, particularly for clothing and furniture. Before 1934, BLS field representatives were instructed to price the “volume selling quality” in each store ([Hoover, 1955](#)). Starting in 1934, the BLS started collecting information on product specifications, in order to attempt to track observationally similar products over time, but even in 1941, a BLS bulletin reports that “the practice is to use the price of the largest selling lines of the current model [...] and to allow the full effect of price changes of the most popular models to the index.” Therefore, quality improvements within product lines could be attributed to price changes rather than quality. In the same bulletin, the BLS acknowledged that this would bias estimates, writing that “when refrigerator prices went down [...] this decline was reflected in the cost-of-living index without adjustment, even though quality had improved so that price, with regard to quality, might have shown a greater decline.” Thus, to the extent that quality improvements were not recognized, we expect the historical index to overstate constant-quality inflation ([U.S. Bureau of Labor Statistics, 1941](#)).

#### 4.1.3 1942–1946

World War II created an unusually severe challenge for the BLS’s fixed-specification approach. Rationing and the diversion of materials to military production made many prewar goods unavailable. The BLS stopped pricing several durable goods, including refrigerators, washing machines, sewing machines, radios, and vacuum cleaners, and adapted its expenditure weights to wartime availability. Where possible, it introduced substitutes: cotton-felt mattresses replaced innerspring mattresses, furniture without spring

filling replaced spring-filled furniture, and rugs with rayon pile replaced all-wool models (U.S. Bureau of Labor Statistics, 1949).

The larger problem for housefurnishings and apparel was that products often deteriorated in quality even when nominal prices remained unchanged. Wartime shortages led to lower wool content, the substitution of rayon for silk, reduced fabric thread counts, and changes in furniture construction. The BLS continued to compare products judged sufficiently similar, but generally lacked a way to value changes in their durability or other characteristics.

The BLS itself emphasized the limits of its measurements. Its postwar assessment estimated that unmeasured deterioration in product quality could have raised the apparel index by another 4–5 percent and the housefurnishings index by another 7–9 percent by the end of 1943. These figures are contemporary estimates, not adjustments actually incorporated in the published wartime indexes. Conversely, as prewar-quality merchandise returned in 1946, quality improvements could go unmeasured in the opposite direction (U.S. Bureau of Labor Statistics, 1949).

#### 4.1.4 1947–1963

After the war, the BLS gradually restored prewar product specifications and expenditure weights, while updating its coverage to reflect the emergence of new consumer goods. In 1947, it expanded the direct pricing of children’s apparel after finding that the prices of children’s clothing did not reliably follow those of comparable adult garments. In 1950, it added televisions to the basket of goods it tracked. The BLS also began to field new consumer expenditure surveys on a more regular basis to update the weights of the CPI (United States Bureau of Labor Statistics, 1966; Rippy, 2014).

#### 4.1.5 1964–1977

This period saw the first major debate over hedonic quality adjustment, largely focused on automobiles. The BLS introduced quality adjustments for new automobiles based on production costs, but did not extend comparable adjustments to other goods (Price Statistics Review Committee, 1961; U.S. Bureau of Labor Statistics, 2020).

#### 4.1.6 1978–1990

The 1978 revision substantially changed how the BLS sampled individual consumer goods, although it retained the basic practice of following a specific item over time. Rather than relying primarily on centrally prescribed product specifications and selected retailers, the BLS used a Point-of-Purchase Survey to identify where households shopped and selected outlets with probabilities related to consumer expenditures. Within each outlet, field representatives selected individual products through a process called *disaggregation*, using relative sales to choose a particular variety, model, or style. They then recorded detailed specifications, including identifying characteristics such as brand and model where applicable, and attempted to reprice the same product in subsequent months (U.S. Bureau of Labor Statistics, 1984; Marcoot, 1985).

#### 4.1.7 1991 onward

Beginning in 1991, the BLS introduced hedonic regression methods to adjust prices for changes in product quality, initially for apparel. These methods estimate the value of observable product characteristics, allowing the BLS to adjust for quality differences when sampled products are replaced. Hedonic adjustments were subsequently introduced for personal computers (1998), televisions (1999), and several major household appliances, including refrigerators, microwave ovens, clothes washers, and dryers (2000), followed by ranges (2008) (U.S. Bureau of Labor Statistics, 2025).

## 4.2 Implementing the historical BLS methodology in Sears data

We next describe how we approximate the historical BLS methodology using our catalog data. During the 1913–1990 period, the BLS primarily measured price changes by repeatedly pricing specific products at sampled retail outlets. When a product disappeared, the BLS selected a replacement and determined whether it was sufficiently comparable to the original product to permit a direct price comparison. If the products were not comparable, the BLS either made an explicit adjustment for differences in quality, linked their price series using observations from a period when both products were available, or imputed the price change using other products in the same category (U.S. Bureau of Labor Statistics, 1984; Moulton and Moses, 1997).

We approximate this approach using three methods to estimate the price at time  $t + 1$  of a product  $i$  observed at time  $t$ :

1. **Matched-model comparison.** If the identical product  $i$  is observed in both periods, we measure its price change directly:

$$\hat{p}_{i,t+1} = p_{i,t+1}. \quad (4.1)$$

2. **Linking using overlapping products.** If product  $i$  disappears, we identify a similar product  $k(i)$  that is observed in both periods  $t$  and  $t + 1$ . Because products  $i$  and  $k$  are both observed in period  $t$ , we can use their overlapping prices to link the two price series. Specifically, we assume that the price of  $i$  would have changed at the same rate as that of  $k$ :

$$\hat{p}_{i,t+1} = p_{i,t} \cdot \frac{p_{k,t+1}}{p_{k,t}}. \quad (4.2)$$

This procedure attributes the difference in the levels of the two products' prices in period  $t$  to differences between the products, rather than treating that difference as inflation. It is an annual-data analogue of the BLS's overlapping-price linking procedure.

3. **Category-level imputation.** If no suitable overlapping product is available, we impute the price change using the observed price movements of other matched products in the same category. Let  $\pi_{j,t+1}$  denote the average price change of products of type  $j$  that can be matched between years  $t$  and  $t + 1$ . We estimate

$$\hat{p}_{i,t+1} = p_{i,t} \cdot \pi_{j,t+1}. \quad (4.3)$$

This assumes that the unobserved price of product  $i$  would have changed at the same rate as the aggregate price of continuing products in its category. The procedure is motivated by the BLS's historical practice of imputing price changes from other valid quotations in the same item category and geographic area.

Implementing these procedures requires identifying products that appear in consecutive Sears catalogs. Unlike modern scanner data, which contain persistent product identifiers, the Sears catalogs do not allow us to directly observe whether a product continues from one year to the next. We therefore infer continuing products using a matching algorithm.

A product in catalog  $t$  and a product in catalog  $t + 1$  are classified as a matched model if they satisfy five criteria: (1) they have the same UNSPSC code; (2) they have the same shipping weight; (3) they have the same brand; (4) they are mutual nearest neighbors in embedding space across the two catalogs; and (5) their cosine similarity exceeds 0.80. Under these criteria, 16% of products have a matched model in the subsequent year. For these products, we can measure annual price changes directly. For products without a matched model, the linking and imputation procedures described above provide estimates of their price changes.

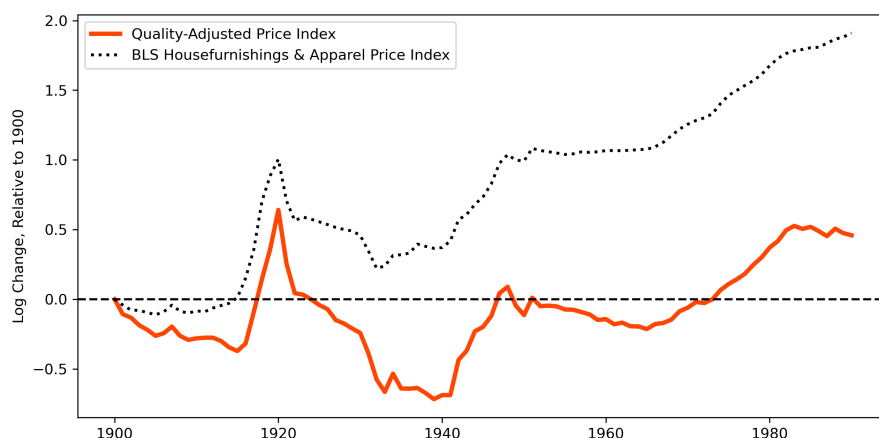
### 4.3 Implementing the modern BLS methodology in the Sears data

The modern BLS methodology, which imputes missing prices with hedonic regressions when identical products cannot be tracked over time, was described in [Section 3.1](#). However, even at the present, the BLS only applies this hedonic methodology to a subset of the products it tracks. The primary reason for this is that hedonic methods require collecting large amounts of detailed attribute data. As a result, the BLS has historically prioritized categories like appliances and apparel, where researchers believed the existing approach was most likely to be biased ([Boskin et al., 1998](#)). In our sample, we extend hedonic methods to a much larger set of goods, including products like vacuum cleaners, mattresses, and furniture. For a smaller subset of goods for which we believe quality adjustments are more limited, like tiepins and bath mats, we instead continue to use the historical methodology. This is purely a practical choice, due to computational and labor intensity of extracting and verifying attributes for each of these product categories. However, our methodology is flexible and could be adapted to incorporate any of these other product categories.

## 5 The cost of living: 1900 to 1990

Using the procedure described in [Section 3.1](#), we construct a quality-adjusted price index for housefurnishings and apparel, shown by the solid red line in [Figure 5](#).<sup>15</sup> Our index declines during the first two decades of the sample, before rising sharply during World War I and falling rapidly in its aftermath. Prices decline further during the Great Depression, then rise again during World War II. In the postwar decades, quality-adjusted prices resume their decline before inflation picks up again in the late 1960s and 1970s.

**Figure 5:** Cost-of-Living Index for Consumer Goods, 1900–1990



The figure shows cumulative log price changes for housefurnishings and apparel, relative to 1900. The solid red line is our quality-adjusted price index, constructed from Sears catalog data using hedonic adjustments for differences in product characteristics. The dotted black line is the historical BLS price index for the corresponding goods, constructed from the sources described in the text. Both series are normalized to zero in 1900.

The dotted black line shows the historical price index for housefurnishings and apparel constructed from official sources.<sup>16</sup> The historical series diverges substantially from our quality-adjusted index,

<sup>15</sup>We use the BLS value for 1907 because we only have a fall catalog for 1907 and a spring catalog for 1906. Goods such as apparel vary substantially across seasons and therefore cannot be compared directly.

<sup>16</sup>For 1900–1913, we use [Rees and Jacobs \(1961\)](#)’s indexes for housefurnishings and apparel. For 1913–1934, 1935–1948,

particularly in the first half of the century. During 1900–1916, 1922–1929, 1932–1939, and 1947–1965, when the historical index shows little inflation on average, our estimates instead imply persistent quality-adjusted deflation.

### 5.1 What accounts for the differences between our index and the BLS?

We next examine how changes in the BLS methodology account for the divergence between the two indexes. [Table 1](#) reports average annual inflation for housefurnishings and apparel under alternative methodologies. We split the data into six periods, corresponding to the six major revisions of the BLS consumer price index that were discussed in [Section 4.1](#).

The first row of [Table 1](#) reports inflation for housefurnishings and apparel based on historical data from the BLS. The next two rows show our replication of the historical BLS matched-model methodology using Sears data, under different assumptions on the accuracy of matching over time. The second row of [Table 1](#) shows estimated annual inflation using what we call a “fuzzy” matched model methodology, meaning that we match products at time  $t$  to the average of their three closest neighbors at time  $t + 1$ . This is meant to capture the fact that historical BLS product descriptions were imprecise, and only allowed BLS data collectors to imperfectly track products over time. The third row shows estimated annual inflation using an exact matched model methodology, as described in [Section 4.2](#). For this methodology, we do our best to track identical products over time in the Sears data. The fuzzy methodology matches the BLS estimates closely for the period of 1900–1941, and the exact methodology matches the BLS well for 1978–1990. For the intermediate period of 1947–1977, the BLS estimates lie in between our fuzzy and exact estimates. These results are consistent with the pattern we documented in [Section 4.1](#), which is that the BLS significantly improved its ability to track products consistently over time throughout this period

**Table 1:** Price Inflation by BLS Methodology Period

| Series                                   | 1900-1913 | 1914-1941 | 1942-1946 | 1947-1963 | 1964-1977 | 1978-1990 |
|------------------------------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1 BLS Apparel + Housefurnishings         | -0.35     | 1.65      | 8.19      | 1.42      | 3.32      | 2.89      |
| 2 Fuzzy Matched Methodology (Sears data) | -0.38     | 1.71      | 11.82     | 1.80      | 4.77      | 4.04      |
| 3 Exact Matched Methodology (Sears data) | -1.40     | -0.01     | 8.29      | 1.12      | 3.02      | 2.91      |
| 4 Modern BLS Methodology (Sears data)    | -2.25     | -0.68     | 10.30     | 0.39      | 3.21      | 2.52      |
| 5 Our Method                             | -2.30     | -1.39     | 11.34     | -0.43     | 2.70      | 2.12      |
| 6 BLS vs. Our Method                     | 1.96      | 3.04      | -3.15     | 1.86      | 0.63      | 0.77      |

*Notes:* The table reports average annual inflation for housefurnishings and apparel, expressed as 100 times the annual change in the log price index, over six periods defined by major changes in BLS methodology. The first row reports inflation from the historical BLS index. The second row approximates the historical BLS matched-model and imputation methodology using Sears catalog data, as described in [Section 4.2](#). The third row applies our approximation to the modern BLS methodology to the Sears data, as described in [Section 4.3](#). The fourth row reports our preferred index, which extends hedonic quality adjustment to additional product categories. The fifth row reports the change in average log posted prices within product categories, without quality adjustment and therefore including changes in product composition. The final row is historical BLS inflation minus inflation under our preferred index, in percentage points per year. The Sears-based indexes use nationally representative BLS expenditure weights.

Rows (4) and (5) illustrate the effects of introducing more systematic quality adjustment. Row (4) applies the modern BLS methodology, described in [Section 4.3](#), to the Sears data. Row (5) extends hedonic quality adjustment to a broader set of furniture and appliances that the BLS does not currently adjust using hedonic methods. Both approaches imply substantially lower inflation than the historical BLS index, with the difference becoming larger as quality adjustment is extended to more goods.

Finally, row (6) reports the annual difference between our preferred index and the historical BLS index. The adjustment is largest before World War II, when our estimates imply inflation that is, on

---

and 1949–1956, we use estimates from BLS historical bulletins nos. 699, 966, and 1256, respectively. From 1957 onward, we use data from the BLS MU series, BLS Bulletin no. 1647, and Federal Reserve Bulletin no. 62.

average, 2.8 percentage points lower per year than the BLS series. After World War II, the difference narrows to 1.1 percentage points annually. The only period in which our index implies *higher* inflation than the BLS is World War II, when severe supply restrictions drastically changed the type of goods supplied. This finding is consistent with contemporaneous BLS reports warning that the CPI likely understated true inflation (U.S. Bureau of Labor Statistics, 1949).

## 5.2 Changes in the quality of goods

The divergence between our index and the historical BLS series is particularly pronounced during periods when the BLS acknowledged difficulties in accounting for changes in product quality. This suggests that improvements in the quality of housefurnishings and apparel may have been an important source of the divergence. We now examine these improvements directly by measuring how the characteristics of goods offered in the Sears catalogs changed over time and estimating their economic value.

For each product category  $j$  and year  $t$ , let  $\Omega_{j,t}$  denote the set of products observed in the catalog, and let  $X_{i,t}$  denote the vector of characteristics of product  $i$ . We define the average characteristics of products in category  $j$  and year  $t$  as

$$\tilde{X}_{j,t} = \frac{1}{|\Omega_{j,t}|} \sum_{i \in \Omega_{j,t}} X_{i,t}.$$

We then estimate the annual change in product quality as

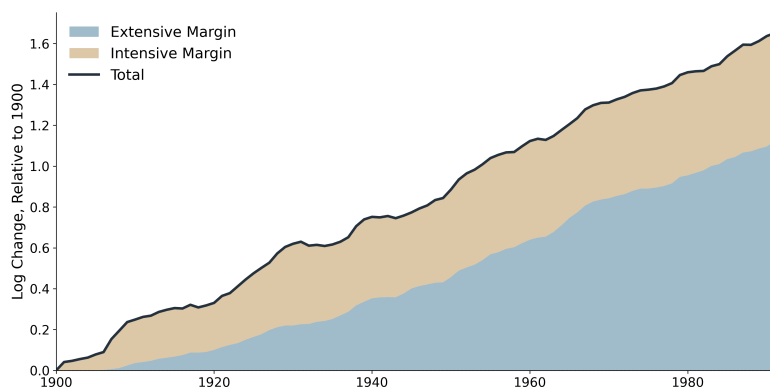
$$q_{j,t} = \left( \tilde{X}_{j,t} - \tilde{X}_{j,t-1} \right)' \hat{\beta}^{j,t}, \quad (5.1)$$

where  $\hat{\beta}^{j,t}$  is the vector of estimated hedonic coefficients from Equation (3.10). These coefficients capture the implicit market valuation of each characteristic, conditional on the other characteristics of the product. Since the hedonic regression uses log prices,  $q_{j,t}$  measures the predicted change in log prices attributable to changes in the characteristics of the average product, holding their implicit valuations fixed at year  $t$ . We show in Section A.4 that under the same assumptions we used to construct the cost-of-living index, Equation (5.1) represents a first order approximation of the willingness to pay for observed changes in product characteristics, as a share of contemporaneous aggregate expenditure.<sup>17</sup> Note that this approach understates quality improvements to the extent that they occur along unobserved dimensions. Our price index methodology tracks the most similar products across years rather than comparing changes in the average product offered, so improvements in unobserved characteristics are effectively differenced out of our price index as well.

We aggregate these annual changes across product categories using the same expenditure weights as in our price index, and plot cumulative quality improvements in Figure 6. We find that quality of goods broadly increased during the twentieth century at a rate of about 1.9% per year. We further decompose these improvements into two components. The *extensive margin* captures the contribution of new features that were not observed in the Sears catalogs during 1900–1905 but appeared in subsequent years. The *intensive margin* captures improvements in characteristics already available at the beginning of the sample, including increases in the prevalence of existing features or the quantity of materials used. We also classify changes in product size, such as increases in television screen size, as intensive margin improvements.

Quality improvement for 1900–1905 is mechanically classified into the intensive margin, by definition. These intensive margin improvements continue to be important over the 20th century as products became bigger. However, the bulk of quality improvement is driven by extensive margin innovations

<sup>17</sup>In practice, we use three-year moving averages of  $\tilde{X}_{j,t}$  to smooth sharp discontinuities in the data. Because some characteristics are missing for some products, we also weight each characteristic's contribution by the share of products for which that characteristic is observed in the base period, which is justified under the assumption in Section A.4.

**Figure 6:** Quality Improvement Along Intensive and Extensive Margins

The figure reports cumulative changes in the estimated economic value of observable product characteristics, in log points relative to 1900. For each characteristic, we multiply the change in its three-year moving average within a product category by its estimated hedonic coefficient and by the share of products reporting that characteristic in the preceding year. The intensive margin includes characteristics with a positive recorded value in 1900–1905 and all size characteristics; the extensive margin includes the remaining characteristics. We first aggregate category-year contributions to BLS categories, then combine them using nationally representative expenditure weights. The shaded regions show the cumulative contributions of these two margins, and the solid line shows their sum.

that were introduced after 1905. These include new materials and technological features which radically transformed the space of consumer goods. This decomposition suggests that about 70% of quality improvement for 1905–1990 was driven by *extensive margin* improvements.

Quality improvements were almost uniformly positive for the period that we study, with one important exception: the Great Depression. Economic theory predicts that during periods of severe economic downturn, households will substitute away from high-quality luxury goods towards lower quality alternatives (Fajgelbaum et al., 2011; Jaimovich et al., 2019). This is exactly what we observe. Between 1931 and 1934, our estimates show the largest sustained quality decline in our sample period, which occurred exclusively along the intensive margin. In Figure A.9, we show some of the margins along which these quality declines were implemented. Panel (a) shows the use of silk in apparel items over the century. Silk was common before the Great Depression, and was used in about half of women’s apparel items in the late 1920s. The use of silk plummeted throughout the entirety of the depression.<sup>18</sup> Panel (b) shows a similar pattern with curtain lengths. Before the Great Depression, curtains were typically 80 to 100 inches long — effectively floor-to-ceiling length. During the depression, average curtain length fell to about 50 inches, just enough to cover the typical US window. Both examples show that the Great Depression appears to have had a large effect on the type of goods that consumers purchased and the rate of quality adjustment.

### 5.2.1 Selected examples of quality change

One key dimension that impacted the quality of goods during this period was the arrival of electricity. In the 1910s, domestic refrigerators were cooled by hand-delivered blocks of ice and washing machines were turned by a hand crank (see Figure A.7); electrification transformed both. Figure A.8 traces how electricity diffused into each. Electricity was adopted more quickly for washing machines and by the 1930s, about half of Sears listed washing machines had an electric motor. It is not until after the war, however, that features common to washing machines today, like water level and temperature control, were introduced. Electric refrigerators appeared later than washing machines but diffused rapidly; by 1939, nearly three-quarters of listed models were electric. Here too, the postwar period brought

<sup>18</sup>This trend is not specific to Sears. [Historical reports](#) on foreign imports show that silk imports fell substantially during the Great Depression and stayed low after World War II

major refinements, including automatic defrosting, water dispensers, and other features. The figure thus highlights substantial quality improvement both before and after World War II, much of which official price statistics may have failed to capture. These quality improvements had important economic consequences; Greenwood et al. (2005) argue that the diffusion of appliances like these, by sharply reducing the time required for laundry, cooking, and cleaning, helped drive the rise in married women's labor force participation over the twentieth century.

**Figure 7:** Women's Dresses, 1922–1928



The figure shows dresses advertised for “stout” women in the Spring 1922, 1925, and 1928 Sears catalogs. Each dress was listed for \$5.98. The images illustrate changes in fabrics and design across similarly priced goods: from gingham in 1922 to silk and wool in 1925 and silk in 1928.

Another example of rapid quality change in our period was women's apparel in the 1920s. Average prices of women's dresses were approximately constant from 1922 to the onset of the Great Depression. Figure 7 shows images of three dresses targeted toward “stout” women, listed for \$5.98 in the spring edition of the 1922, 1925, and 1928 Sears catalogs. In 1922, \$5.98 could buy a plaid gingham dress that

emphasized its good quality and durability. By 1925, the same amount could garner a more intricate silk and wool dress with embroidered sleeves. The equivalently priced dress in 1928 was made entirely of silk and had an even more ornate cut and pattern. Materials improvements, for instance with the proliferation of silk, explain a large portion of quality improvement in women’s dresses in the 1920s, but changes in style and design intricacy are also significant. Therefore, our quality-adjusted price index finds that quality-adjusted prices fell substantially over this period, even if the average sticker price remained constant.

### 5.3 Robustness checks

#### 5.3.1 Robustness to Changing Characteristics

A key advantage of a hedonic regression, relative to looking only at price changes of the subset of goods which appear in consecutive years, is that we can study implicit price dynamics of *all goods*, not just surviving goods. However, the framework we presented in [Section 3](#) implicitly assumes that even if the set of goods available changes from one year to the next, the underlying set of characteristics of the goods available at each time  $t - 1$  continue to exist at time  $t$ , such that the time  $t$  hedonic schedule can be used to predict prices for goods in the previous year. This assumption is, of course, imperfect. There are many characteristics that disappeared at discrete points in our sample to make way for new product characteristics, from wringers on washing machines to girdles on dresses.

As a robustness exercise, we compute the cost-of-living index in [Equation \(3.12\)](#), excluding products that are very different from all of the products in the following year (i.e. exiting goods). Empirically, we implement this using text embeddings,  $\phi_i$ . For any pair of goods in category  $j$ , denoted  $i$  and  $i'$ , the cosine similarity of the product is given by,

$$\text{sim}(i, i') = \frac{\phi_i \cdot \phi_{i'}}{\|\phi_i\| \|\phi_{i'}\|}$$

For each product  $i$  at time  $t$ , we construct a measure of its novelty relative to time  $t + 1$  as,

$$\sigma_{i,t+1} = \max_{i' \in \mathcal{I}_{j,t+1}} \text{sim}(i, i')$$

The novelty variable  $\sigma_{i,t+1}$  captures how similar product  $i$  is to its most similar alternative at time  $t + 1$ . Similarly, we can compute  $\sigma_{i,t-1}$  as similarity to products in the previous period. We then compute the cost-of-living index, excluding all products for which  $\sigma_{i,t+1} < 0.7$ . This excludes 11.1% of products which were significantly different from any product in the next year. We also compute a version in which we exclude all products for which  $\sigma_{i,t-1} < 0.7$  or  $\sigma_{i,t+1} < 0.7$ , such that only “stable” goods are included. The new cost-of-living index using each of these samples is shown in the solid grey and red lines, respectively, in [Figure A.4](#). Our baseline estimates are shown in the dashed line. The three series are very similar over the entire period, suggesting that our main conclusions are not driven by new or dying goods.

#### 5.3.2 Robustness to Sears Sample

To construct a cost-of-living index, we need a representative sample of all goods and services purchased by Americans in each year. Sears was clearly not representative of many sectors of the economy, including services, food and automobiles, but our analysis thus far has assumed that it was broadly representative of housefurnishings and apparel purchased by Americans. According to an [internal document](#) released in 1978, Sears described itself as follows: “We are not a fashion store; we are not a store for the whimsical nor the affluent. Sears is a family store for middleclass, homeownership Americans.” Moreover, Sears’s mail-order business made its products accessible to effectively all households throughout the

United States, including in rural areas (Raff, 2006). By competing directly with local retailers and other mail-order firms (Michael, 1994), Sears operated in a relatively integrated national retail market. Assuming relatively stable markups across retailers, competitive pressures would imply that long-run inflation at Sears should broadly track inflation at other retail outlets.

To assess the representativeness of Sears, we supplement our data collection with J.C. Penney catalogs for the period of 1969–1982. J.C. Penney targeted a somewhat higher end of the market, and therefore will give us a sense of whether different tiers of the retail sectors saw very different movements in prices. We construct quality-adjusted price indexes following the methodology in Section 3.1 for both J.C. Penney and Sears for this period. We restrict to the sample of products for which the two catalogs overlap, since J.C. Penney did not sell some major appliances such as refrigerators and ranges. For the period of overlap, the indexes, show in Figure A.5, are fairly consistent, indicating high inflation in the 70s.

As a further test of external validity, we examine whether the hedonic model estimated using Sears products generalizes to J.C. Penney. For each J.C. Penney product, we predict its price using both the Sears hedonic regression from Equation (3.10) and a separate hedonic regression estimated on J.C. Penney data. Figure A.6 compares the two sets of predicted prices, which are remarkably similar. Excluding a small fraction of outliers, the  $R^2$  between the two predictions is 0.90. This strong agreement suggests that the implicit valuations of product characteristics estimated from Sears data generalize to other retailers, providing further support for the external validity of our hedonic estimates.

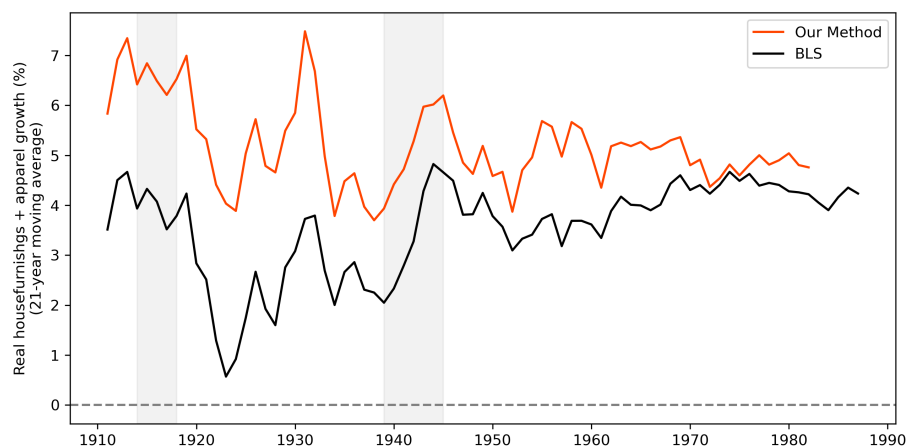
## 6 Implications for real growth

Our revisions to the price index for housefurnishings and apparel have direct implications for measured real consumption growth, both within these sectors and for aggregate consumption. Table 2 reports annualized real expenditure growth under our deflator and the corresponding BLS deflator at three levels of aggregation. The first two columns focus on housefurnishings and apparel, holding nominal expenditures fixed and varying only the price index used to deflate them. A moving average of growth in these sectors is also shown in Figure 8. Our methodology implies higher real consumption growth in nearly every period, with the largest differences occurring early in the twentieth century. Moreover, our estimates reveal a pronounced slowdown in the growth of real consumption in these sectors over the course of the century. This pattern is consistent with the hypothesis that technological progress becomes increasingly difficult to sustain as the stock of existing knowledge expands (Bloom et al., 2020).

To examine the implications for aggregate goods consumption, we combine our estimates with expenditure and price data for other major consumption categories. We apply our quality-adjusted price index to apparel, housefurnishings, entertainment and recreational goods.<sup>19</sup> For the remaining sectors, we use independent price indexes. Specifically, we distinguish food, automobiles, fuel, and other nondurable goods, including tobacco, pharmaceuticals, personal care products, and household supplies. We obtain nominal expenditures from historical reconstructions of personal consumption expenditures before 1929 and from the Bureau of Economic Analysis (BEA) thereafter.<sup>20</sup> Importantly, we also incorporate independent estimates of automobile quality improvements, which were substantial during the first half of the twentieth century. For 1906–1940, we use the hedonic automobile price index constructed by Raff and Trajtenberg (1997), and for 1948–1965 we use the quality-adjusted index

<sup>19</sup>We include in this group miscellaneous durable goods, which includes luggage, telephones, therapeutic appliances and jewelry. We assume these goods have similar quality-adjusted price dynamics to the housefurnishings and apparel sector

<sup>20</sup>For 1900–1928, nominal expenditures are drawn from Stanley Lebergott’s historical reconstruction, reported in *Historical Statistics of the United States*, series Cd1–77. From 1929 onward, we use BEA personal consumption expenditure data. For food prices, we use the historical BLS food CPI before 1929 and the BEA food deflator thereafter. For fuel, we use the NBER historical wholesale price index for fuel and lighting before 1929, spliced to the BEA gasoline and energy goods deflator. For other nondurable goods, we construct an implicit price deflator from Lebergott’s nominal and constant-dollar expenditure series before 1929, and aggregate detailed BEA price indexes for the corresponding categories thereafter.

**Figure 8:** Real Growth of Housefurnishings & Apparel

The figure plots 21-year centered moving averages of annual percentage growth in real U.S. expenditures on housefurnishings and apparel. We deflate the same nominal expenditures using either our quality-adjusted Sears price index (red line) or the historical BLS benchmark (black line). Annual growth is calculated from percentage changes in real expenditure levels, without a per-capita adjustment. The shaded areas mark World War I (1914–1918) and World War II (1939–1945).

constructed by [Gordon \(1990\)](#), both reported in *Historical Statistics of the United States* (series Cc63–64). We use the BEA automobile price index for the intervening years and from 1966 onward.

For apparel, housefurnishings, entertainment and recreational goods, we alternatively apply our price index or the corresponding BLS index, holding the deflators for all remaining sectors fixed. We aggregate across sectors using an annually chained Fisher price index, calculated as the geometric mean of Laspeyres and Paasche price relatives, following the methodology of the Bureau of Economic Analysis (BEA). To maintain consistent sector definitions throughout the twentieth century, however, we use broader sectoral categories than those employed in modern national accounts. We then deflate total nominal goods expenditures by the resulting aggregate price index. The middle two columns of [Table 2](#) report the implied growth in real goods consumption under the two methodologies. Despite leaving the price indexes for food, automobiles, fuel, and other nondurable goods unchanged, our quality adjustment implies substantially higher real goods consumption growth, particularly in the first half of the twentieth century.

Finally, we extend the analysis to total personal consumption expenditures by incorporating services and housing. We separate expenditures on nonhousing services from actual and imputed housing rents, and construct independent price indexes for each. For nonhousing services, we recover an implicit price index for 1900–1928 using Lebergott’s historical estimates of nominal and constant-dollar expenditures. From 1929 onward, we use published BEA price indexes for services excluding energy and housing and for electricity and gas services, which we combine using a chain-type Fisher index with annual expenditure weights. We splice the historical series to the BEA index in 1929. For housing, we deflate both actual and imputed rents using the national rental price index constructed by [Lyons et al. \(2026\)](#), which uses historical real estate listings to track quality-adjusted rents over the twentieth century.<sup>21</sup>

We aggregate the resulting seven expenditure categories using the same chained Fisher methodology. The final two columns of [Table 2](#) show that our revision raises measured annual growth in total real consumption expenditures by approximately 0.6 percentage points during 1900–1941 and 0.3 percentage points during 1946–1990. Thus, even when incorporated into an aggregate that includes all consumer

<sup>21</sup> Historical services prices are considerably less well measured than goods prices. Unlike goods, for which relatively detailed price quotations are often available, many services lack consistent historical price series, and changes in their quality are particularly difficult to measure. Since we hold the services and housing deflators fixed across our two specifications, these measurement limitations do not directly generate the differences between our method and the BLS benchmark.

goods, services, and housing, our revised estimates of quality-adjusted prices imply a substantial increase in measured consumption growth over the twentieth century.

**Table 2:** Real Growth Over the Twentieth Century

| Period    | Housefurnishings<br>+ apparel |      | All consumption<br>goods |      | Total consumption<br>expenditures |      |
|-----------|-------------------------------|------|--------------------------|------|-----------------------------------|------|
|           | Our method                    | BLS  | Our method               | BLS  | Our method                        | BLS  |
|           | (1)                           | (2)  | (3)                      | (4)  | (5)                               | (6)  |
| 1900–1913 | 8.46                          | 6.36 | 5.09                     | 4.50 | 4.25                              | 3.89 |
| 1914–1928 | 6.53                          | 3.85 | 3.13                     | 2.23 | 2.78                              | 2.27 |
| 1929–1950 | 3.89                          | 1.90 | 3.23                     | 2.55 | 2.69                              | 2.32 |
| 1951–1964 | 5.10                          | 3.52 | 4.11                     | 3.57 | 4.29                              | 3.99 |
| 1965–1972 | 4.63                          | 4.04 | 4.20                     | 3.98 | 3.95                              | 3.84 |
| 1973–1979 | 3.77                          | 4.08 | 2.28                     | 2.39 | 2.72                              | 2.77 |
| 1980–1990 | 5.94                          | 4.41 | 3.89                     | 3.31 | 3.87                              | 3.64 |
| 1900–1941 | 6.12                          | 3.29 | 3.74                     | 2.82 | 3.08                              | 2.56 |
| 1946–1990 | 4.81                          | 3.62 | 3.65                     | 3.22 | 3.75                              | 3.53 |
| 1900–1990 | 5.28                          | 3.59 | 3.64                     | 3.08 | 3.41                              | 3.10 |

The table reports compound annual growth rates (percent) in real consumption expenditures between the endpoints of each period under our quality-adjusted price index and the historical BLS benchmark. The first two columns report housefurnishings and apparel, holding nominal expenditures fixed and varying the price index. The next two columns report all consumption goods, aggregated across five sectors: food, automobiles, fuel, housefurnishings and apparel together with other recreational and durable goods, and remaining nondurable goods. The final two columns report total consumption expenditures, adding nonhousing services and housing rents to the same five goods sectors. Nominal expenditures are taken from historical national accounts before 1929 and BEA data thereafter. The five-sector goods and seven-sector total-consumption price indexes use annually chained Fisher links, calculated as the geometric mean of Laspeyres and Paasche price relatives with preceding- and current-year expenditure shares, respectively. For nonhousing services, we splice the historical implicit price index to a BEA component-based index in 1929; rents are deflated separately. The only difference between our method and the BLS benchmark is the price index applied to the housefurnishings, apparel, and additional quality-sensitive goods sector; all other sector deflators are held fixed.

## 6.1 Total Factor Productivity

The timing of peak productivity growth in the United States has been the subject of considerable debate. [Gordon \(1999\)](#) characterized twentieth-century productivity growth as “one big wave,” with rapid growth from the early twentieth century through the 1960s and a peak in 1950–1964. Subsequent revisions to the national accounts, including the introduction of chain-weighted real GDP measures, substantially raised estimated growth during 1928–1950, leading Gordon’s later estimates to identify this earlier period as the peak ([Gordon, 2013, 2016](#)). [Field \(2003\)](#) argued that the 1930s were the most technologically progressive period of the century. Using the historical productivity estimates of [Kendrick \(1961\)](#) and extending the endpoint of the Depression-era comparison from 1937 to 1941, Field found exceptionally rapid TFP growth in 1929–1941, attributing much of the subsequent postwar productivity gains to technological and organizational improvements achieved before World War II. More recently, [Field \(2023\)](#) argued that manufacturing TFP actually declined during the war, challenging accounts that emphasize wartime learning-by-doing. In contrast, [Bakker et al. \(2019\)](#) reconstructed prewar growth accounts using more comprehensive industry coverage, improved measures of labor quality, and capital-services-based inputs. These adjustments reduce estimated prewar TFP growth relative to Kendrick’s estimates. They also question the use of annually chain-linked output measures during the volatile 1930s, arguing that these may suffer from chain drift. Their preferred estimates imply annual TFP growth of 1.86% in 1929–1941, compared with 1.98% in 1948–1960 and 2.21% in 1960–1973, using postwar BLS estimates for the latter periods. [Crafts and Woltjer \(2021\)](#) reaffirm this chronology, placing the peak of U.S. TFP growth in the 1960s and early 1970s.

**Table 3:** Original and Adjusted TFP Growth

| Period    | TFP growth (% per annum) |              |
|-----------|--------------------------|--------------|
|           | Original estimate        | New estimate |
| 1899–1909 | 0.93                     | 1.30         |
| 1909–1919 | 0.64                     | 1.25         |
| 1919–1929 | 1.63                     | 2.00         |
| 1929–1941 | 1.86                     | 2.49         |
| 1948–1960 | 1.98                     | 2.43         |
| 1960–1973 | 2.21                     | 2.38         |
| 1973–1989 | 0.48                     | 0.55         |

The table reports original average annual TFP growth estimates (percent) from Bakker et al. (2019), Table 4, together with estimates adjusted for the difference between our total-consumption deflator and the BLS-based benchmark. For each period, we calculate the average annual difference between the log price growth of total consumption under our method and under the BLS benchmark, multiply by 100, and subtract it from the original TFP growth rate. Annual price changes are assigned to the period ending in the indicated year, so that a period beginning in year  $a$  uses price changes from  $a + 1$  onward. This adjustment assumes that bias in the private domestic economy deflator is equal to bias in our total-consumption deflator, holding all other growth-accounting components fixed. We assume that the bias for the year between 1899 and 1900 is the same as for the period of 1900–1909, since our sample begins in 1900.

To construct a revised TFP series, we begin with the estimates of Bakker et al. (2019) and adjust them using the price index bias estimated in Section 5. Specifically, we assume that the bias in the price deflator for the private domestic economy (PDE) is equal to our estimated bias for total consumption expenditures. This assumption is plausible given that national accounts rely on many of the same underlying price data as BLS consumer price indexes. Moreover, Gordon (1990) documents substantial quality-adjustment bias in investment goods, suggesting that the bias for the broader economy could be even larger than that observed in our consumer goods sample. Under this assumption, we adjust the TFP growth rates reported by Bakker et al. (2019) by subtracting the estimated deflator bias, expressed in log terms, while holding their other growth-accounting assumptions fixed. The original and adjusted TFP estimates are shown in Table 3.

These adjustments shift the estimated timing of peak productivity growth earlier in the twentieth century. Whereas Bakker et al. (2019) find that TFP growth peaked in the 1960s and early 1970s, our adjusted estimates imply slightly higher growth in 1929–1941 and 1948–1960, although the differences across these periods are small. This pattern is consistent with Gordon’s characterization of “one big wave” of productivity growth centered around World War II. Consistent with Field (2003), we find rapid productivity growth during the Great Depression, but our estimates suggest that this wave began considerably earlier: annual TFP growth exceeds 1% in every period except 1973–1989. The resulting slowdown in the final decades of the century is particularly striking, with TFP growth in 1973–1989 falling to less than half the rate of any preceding period.

## 7 Conclusion

This paper constructs a quality-adjusted price index for U.S. apparel and housefurnishings spanning the twentieth century. Using newly digitized catalog data, we recover millions of product-level prices and descriptions, estimate hedonic price schedules from detailed product characteristics, and use those schedules to price fixed bundles of product characteristics over time. The resulting index offers a new view of long-run inflation, quality change, and real growth in the retail goods sector.

Our central finding is that quality improvements in housefurnishings and apparel over the twentieth century were substantial, and that existing price indexes failed to fully account for them. Our revision

reflects two important differences in methodology. First, we use detailed product descriptions to identify identical or highly similar goods across adjacent years, allowing us to track their prices more accurately. Second, when products cannot be matched, we use hedonic methods to impute price changes that account for differences in product characteristics. While previous research has emphasized the importance of hedonic quality adjustment, we show that improved product matching is also quantitatively important, particularly before 1978, when historical BLS records indicate that the agency often struggled to track identical products over time. The magnitudes are large: replacing standard deflators with our index raises measured growth in real consumption of the sectors we study from a factor of 24 to 102 between 1900 and 1990, and raises aggregate real expenditure growth from 15.6 to 20.4 once other sectors are held fixed.

The timing of this revision is also important. Conventional estimates imply that goods consumption accelerated after World War II. Our quality-adjusted index reverses this pattern. Real consumption growth of housefurnishings and apparel was faster from 1900 to 1941 than from 1946 to 1990. This suggests that a substantial share of twentieth-century growth occurred before the war, during precisely the period for which historical price measurement is weakest.

Our index covers only consumer goods and omits major sectors such as services, food, and housing. Even holding inflation in those sectors fixed, quality adjustment of consumer goods alone materially raises measured real expenditure growth over the century. More broadly, scalable methods for digitizing and analyzing historical text open new opportunities to measure economic change over periods for which structured data have never existed. Extending this approach to other retailers, countries, and historical records is a natural next step toward a fuller account of one of the most transformative periods in economic history.

## References

- Moses Abramovitz and Paul A. David. American macroeconomic growth in the era of knowledge-based progress: The long-run perspective. In Stanley L. Engerman and Robert E. Gallman, editors, *The Cambridge Economic History of the United States*, volume 3, pages 1–92. Cambridge University Press, Cambridge, 2000. doi: 10.1017/CHOL9780521553087.002.
- Philippe Aghion, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li. Missing growth from creative destruction. *American Economic Review*, 109(8):2795–2822, 2019.
- Hemanth Asirvatham, Elliott Mokski, and Andrei Shleifer. GPT as a measurement tool. NBER Working Paper 34834, National Bureau of Economic Research, February 2026. URL <https://doi.org/10.3386/w34834>.
- Verónica Bäcker-Peral, Vitaly Meursault, and Christopher Severen. Can LLMs credibly transform the creation of panel data from diverse historical tables? Working Paper 25-28, Federal Reserve Bank of Philadelphia, September 2025.
- Gerben Bakker, Nicholas Crafts, and Pieter Woltjer. The sources of growth in a technologically progressive economy: The United States, 1899–1941. *The Economic Journal*, 129(622):2267–2294, 2019. doi: 10.1093/ej/uez002.
- Mark Bills. Do higher prices for new goods reflect quality growth or inflation? *Quarterly Journal of Economics*, 124(2):637–675, 2009.
- Mark Bills and Peter J. Klenow. Quantifying quality growth. *American Economic Review*, 91(4): 1006–1030, 2001. doi: 10.1257/aer.91.4.1006.
- Nicholas Bloom, Charles I. Jones, John Van Reenen, and Michael Webb. Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144, 2020. doi: 10.1257/aer.20180338.
- Michael J. Boskin, Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale W. Jorgenson. Consumer prices, the Consumer Price Index, and the cost of living. *Journal of Economic Perspectives*, 12(1):3–26, 1998. doi: 10.1257/jep.12.1.3.
- Christian Broda and David E. Weinstein. Globalization and the gains from variety. *The Quarterly Journal of Economics*, 121(2):541–585, 2006. doi: 10.1162/qjec.2006.121.2.541.
- Dora L. Costa. Estimating real income in the United States from 1888 to 1994: Correcting CPI bias using Engel curves. *Journal of Political Economy*, 109(6):1288–1310, 2001. doi: 10.1086/323279.
- Andrew T. Court. Hedonic price indexes with automotive examples. In *The Dynamics of Automobile Demand*, pages 98–119. General Motors Corporation, New York, 1939.
- Nicholas Crafts and Pieter Woltjer. Growth accounting in economic history: Findings, lessons and new directions. *Journal of Economic Surveys*, 35(3):670–696, 2021. doi: 10.1111/joes.12348.
- W. Erwin Diewert. Exact and superlative index numbers. *Journal of Econometrics*, 4(2):115–145, 1976. doi: 10.1016/0304-4076(76)90009-9.
- Gabriel Ehrlich, John Haltiwanger, Ron Jarmin, David Johnson, Ed Olivares, Luke Pardue, Matthew D. Shapiro, and Laura Yi Zhao. Quality adjustment at scale: Hedonic versus exact demand-based price indices. *American Economic Review*, 116(6):1955–1995, 2026. doi: 10.1257/aer.20230766.

- Pablo D. Fajgelbaum, Gene M. Grossman, and Elhanan Helpman. Income distribution, product quality, and international trade. *Journal of Political Economy*, 119(4):721–765, 2011. doi: 10.1086/662628.
- Robert C. Feenstra. New product varieties and the measurement of international prices. *American Economic Review*, 84(1):157–177, 1994.
- Alexander J. Field. The most technologically progressive decade of the century. *American Economic Review*, 93(4):1399–1413, 2003. doi: 10.1257/000282803769206377.
- Alexander J. Field. The decline of US manufacturing productivity between 1941 and 1948. *The Economic History Review*, 76(4):1163–1190, 2023. doi: 10.1111/ehr.13239.
- Robert J. Gordon. *The Measurement of Durable Goods Prices*. National Bureau of Economic Research Monograph. University of Chicago Press, Chicago, 1990.
- Robert J. Gordon. U.S. economic growth since 1870: One big wave? *American Economic Review*, 89(2): 123–128, 1999. doi: 10.1257/aer.89.2.123.
- Robert J. Gordon. Revisiting U.S. productivity growth over the past century with a view of the future. In D. S. Prasada Rao and Bart van Ark, editors, *World Economic Performance: Past, Present and Future*, chapter 12, pages 317–370. Edward Elgar Publishing, 2013. doi: 10.4337/9781781953556.00017.
- Robert J. Gordon. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton University Press, Princeton, NJ, 2016. ISBN 9780691147727.
- Jeremy Greenwood, Ananth Seshadri, and Mehmet Yorukoglu. Engines of liberation. *The Review of Economic Studies*, 72(1):109–133, 2005. doi: 10.1111/0034-6527.00326.
- Zvi Griliches. Hedonic price indexes for automobiles: An econometric analysis of quality change. In *The Price Statistics of the Federal Government*, pages 173–196. National Bureau of Economic Research, New York, 1961. Report of the Price Statistics Review Committee.
- Zvi Griliches, editor. *Price Indexes and Quality Change: Studies in New Methods of Measurement*. Harvard University Press, Cambridge, MA, 1971.
- Bruce W. Hamilton. Using Engel’s Law to estimate CPI bias. *American Economic Review*, 91(3): 619–630, 2001. doi: 10.1257/aer.91.3.619.
- Jerry A. Hausman. Cellular telephone, new products and the CPI. SSRN Scholarly Paper, March 1997. URL <https://ssrn.com/abstract=225761>.
- Ethel D. Hoover. Average retail prices: Collection and calculation techniques and problems. Bulletin 1182, U.S. Bureau of Labor Statistics, Washington, DC, June 1955. URL [https://fraser.stlouisfed.org/files/docs/publications/bls/bls\\_1182\\_1955.pdf](https://fraser.stlouisfed.org/files/docs/publications/bls/bls_1182_1955.pdf).
- Nir Jaimovich, Sergio Rebelo, and Arlene Wong. Trading down and the business cycle. *Journal of Monetary Economics*, 102:96–121, 2019.
- Òscar Jordà, Moritz Schularick, and Alan M. Taylor. Macrofinancial history and the new business cycle facts. *NBER Macroeconomics Annual*, 31(1):213–263, 2017. doi: 10.1086/690241.
- John W. Kendrick. *Productivity Trends in the United States*. Number 71 in General Series. Princeton University Press, Princeton, NJ, 1961. Published for the National Bureau of Economic Research.
- A. A. Konüs. The problem of the true index of the cost of living. *Econometrica*, 7(1):10–29, 1939. Translated from the 1924 Russian original.

- Ronan C. Lyons, Allison Shertzer, Rowena Gray, and David Agorastos. The price of housing in the United States, 1890–2006. *The Quarterly Journal of Economics*, 141(1):559–603, 2026. doi: 10.1093/qje/qjaf047.
- John L. Marcoot. Revision of consumer price index is now under way. *Monthly Labor Review*, pages 27–38, April 1985. URL <https://www.bls.gov/opub/mlr/1985/04/art3full.pdf>.
- Steven C. Michael. Competition in organizational form: Mail order versus retail stores, 1910–1940. *Journal of Economic Behavior & Organization*, 23(3):269–286, 1994. doi: 10.1016/0167-2681(94)90002-7.
- Brent R. Moulton and Karin E. Moses. Addressing the quality change issue in the consumer price index. *Brookings Papers on Economic Activity*, (1):305–366, 1997. URL [https://www.brookings.edu/wp-content/uploads/1997/01/1997a\\_bpea\\_moulton\\_moses\\_gordon\\_bosworth.pdf](https://www.brookings.edu/wp-content/uploads/1997/01/1997a_bpea_moulton_moses_gordon_bosworth.pdf).
- Emi Nakamura and Jón Steinsson. Lost in transit: Product replacement bias and pricing to market. *American Economic Review*, 102(7):3277–3316, 2012. doi: 10.1257/aer.102.7.3277.
- William D. Nordhaus. Do real-output and real-wage measures capture reality? The history of lighting suggests not. In Timothy F. Bresnahan and Robert J. Gordon, editors, *The Economics of New Goods*, pages 27–70. University of Chicago Press, Chicago, 1996.
- Ariel Pakes. A reconsideration of hedonic price indexes with an application to PC’s. *American Economic Review*, 93(5):1578–1596, 2003. doi: 10.1257/000282803322655455.
- Price Statistics Review Committee. *The Price Statistics of the Federal Government: Review, Appraisal, and Recommendations*. Number 73 in General Series. National Bureau of Economic Research, 1961. URL <https://www.nber.org/books-and-chapters/price-statistics-federal-government>.
- Daniel M. G. Raff. Distribution. In Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, editors, *Historical Statistics of the United States: Millennial Edition*, volume 4, pages 705–712. Cambridge University Press, 2006. URL <https://hsus.cambridge.org/HSUSWeb/essay/pdf/De.ESS.01.pdf>.
- Daniel M. G. Raff and Manuel Trajtenberg. Quality-adjusted prices for the american automobile industry: 1906–1940. In Timothy F. Bresnahan and Robert J. Gordon, editors, *The Economics of New Goods*, volume 58 of *Studies in Income and Wealth*, pages 71–108. University of Chicago Press, Chicago, 1997. URL <https://www.nber.org/chapters/c6065>.
- Stephen J. Redding and David E. Weinstein. Measuring aggregate price indices with taste shocks: Theory and evidence for CES preferences. *The Quarterly Journal of Economics*, 135(1):503–560, 2020. doi: 10.1093/qje/qjz031.
- Stephen B. Reed. One hundred years of price change: The Consumer Price Index and the American inflation experience. *Monthly Labor Review*, 2014.
- Albert Rees and Donald P. Jacobs. The cost-of-living index. In *Real Wages in Manufacturing, 1890–1914*, pages 74–119. Princeton University Press, 1961.
- Darren Rippy. The first hundred years of the consumer price index: A methodological and political history. *Monthly Labor Review*, April 2014. doi: 10.21916/mlr.2014.13. URL <https://www.bls.gov/opub/mlr/2014/article/the-first-hundred-years-of-the-consumer-price-index.htm>.
- Sherwin Rosen. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55, 1974. doi: 10.1086/260169.

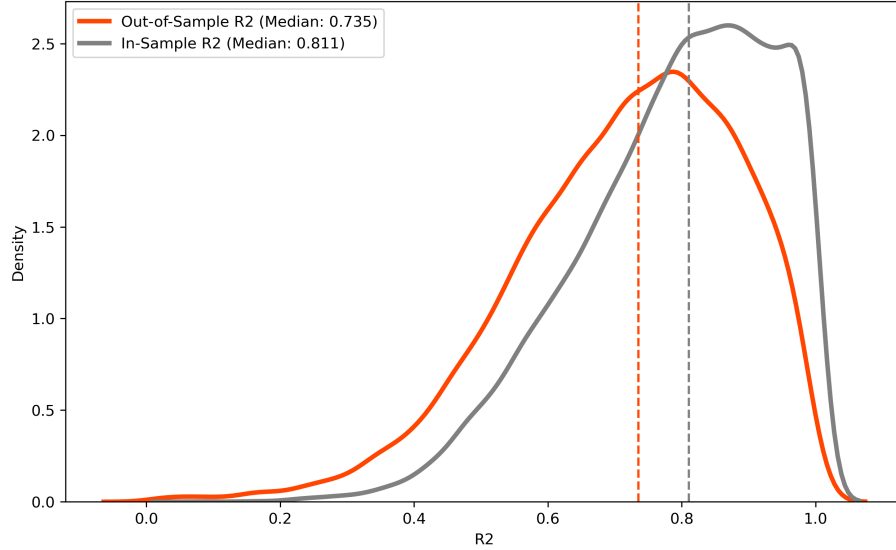
- Kazuo Sato. The ideal log-change index number. *The Review of Economics and Statistics*, 58(2):223–228, 1976. doi: 10.2307/1924029.
- Jack E. Triplett. The economic interpretation of hedonic methods. *Survey of Current Business*, 66(1): 36–40, 1986.
- Jack E. Triplett. *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*. OECD Publishing, Paris, 2006. doi: 10.1787/9789264028159-en.
- United States Bureau of Labor Statistics. Index numbers of wholesale prices in the United States and foreign countries. Technical Report 173, United States Bureau of Labor Statistics, Washington, DC, July 1915. URL <https://fraser.stlouisfed.org/title/index-numbers-wholesale-prices-united-states-foreign-countries-3828>. Wholesale Prices Series, No. 3.
- United States Bureau of Labor Statistics. The Consumer Price Index: History and techniques. Bulletin of the United States Bureau of Labor Statistics 1517, United States Department of Labor, Bureau of Labor Statistics, Washington, DC, 1966.
- U.S. Bureau of Labor Statistics. R-CPI-U-RS changes over time. Bureau of Labor Statistics website. URL <https://www.bls.gov/cpi/research-series/r-cpi-u-rs-changes.htm>. Accessed October 2026.
- U.S. Bureau of Labor Statistics. Methods of procuring and computing statistical information of the bureau of labor statistics. Bulletin 326, U.S. Department of Labor, Bureau of Labor Statistics, Washington, DC, March 1923. URL <https://fraser.stlouisfed.org/title/methods-procuring-computing-statistical-information-bureau-labor-statistics-3982/fulltext>.
- U.S. Bureau of Labor Statistics. Changes in cost of living in large cities in the united states, 1913–41. Bulletin 699, U.S. Department of Labor, Bureau of Labor Statistics, Washington, DC, 1941.
- U.S. Bureau of Labor Statistics. Consumers’ prices in the united states, 1942–48: Analysis of changes in cost of living. Bulletin 966, U.S. Bureau of Labor Statistics, 1949. URL <https://fraser.stlouisfed.org/title/consumers-prices-united-states-1942-48-analysis-changes-cost-living-4382/fulltext>.
- U.S. Bureau of Labor Statistics. BLS handbook of methods, volume ii: The consumer price index. Bulletin 2134-2, U.S. Bureau of Labor Statistics, April 1984. URL <https://fraser.stlouisfed.org/title/bls-handbook-methods-5238/bls-handbook-methods-consumer-price-index-volume-ii-529770/fulltext>.
- U.S. Bureau of Labor Statistics. The consumer price index. Technical report, U.S. Bureau of Labor Statistics, 2018. URL <https://www.bls.gov/opub/hom/pdf/cpi-20180214.pdf>. Chapter 17, updated February 14, 2018, p. 2.
- U.S. Bureau of Labor Statistics. Chronology of changes in the consumer price index. Bureau of Labor Statistics website, April 2020. URL <https://www.bls.gov/cpi/additional-resources/historical-changes.htm>.
- U.S. Bureau of Labor Statistics. Quality adjustment in the cpi: Ranges and cooktops, 2025. URL <https://www.bls.gov/cpi/quality-adjustment/ranges.htm>. Last modified July 25, 2025.

Yrjö O. Vartia. *Relative Changes and Index Numbers*. Number 4 in ETLA Series A. The Research Institute of the Finnish Economy, Helsinki, 1976.

## A Appendix

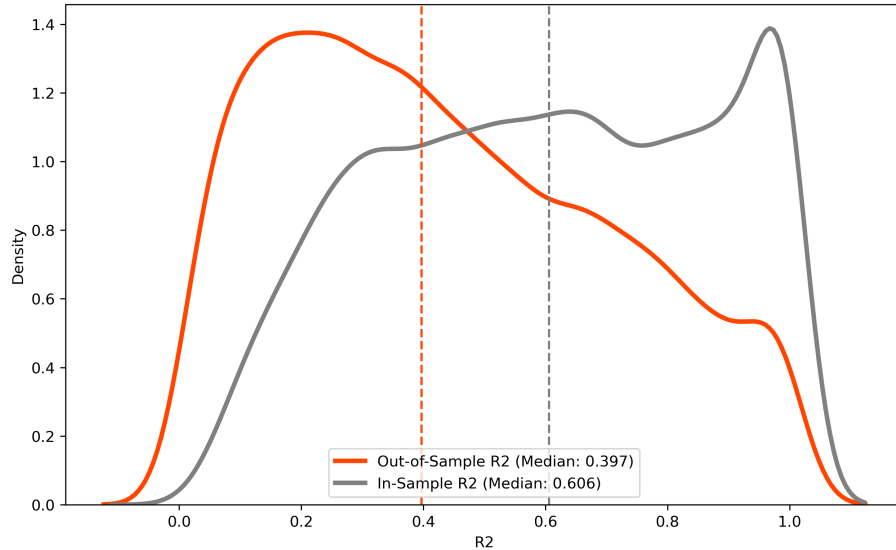
### A.1 Figures

**Figure A.1:** Distribution of  $R^2$  For Hedonic Regressions

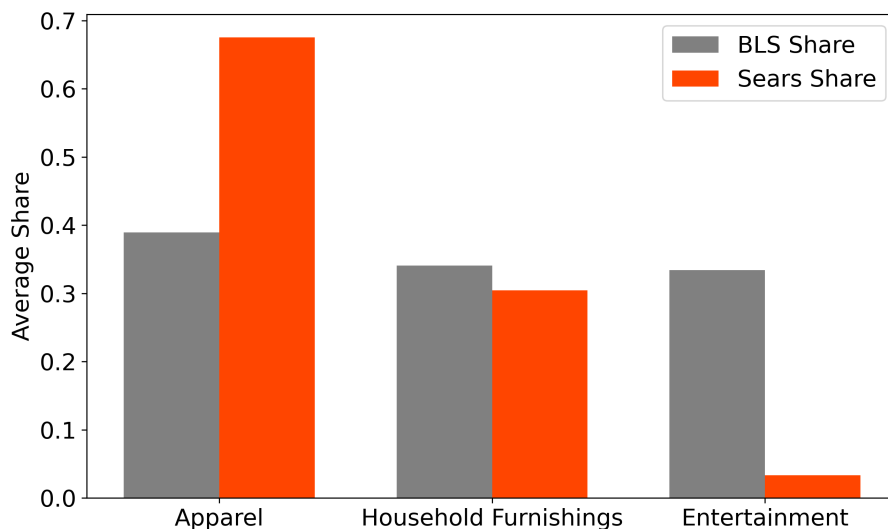


The figure shows kernel density estimates of in-sample and out-of-sample  $R^2$  from the ridge hedonic regressions of log prices on product characteristics, estimated separately by product category and year. Both distributions are restricted to regressions with positive out-of-sample  $R^2$ . The dashed vertical lines indicate the medians for this restricted sample, which are also reported in the legend.

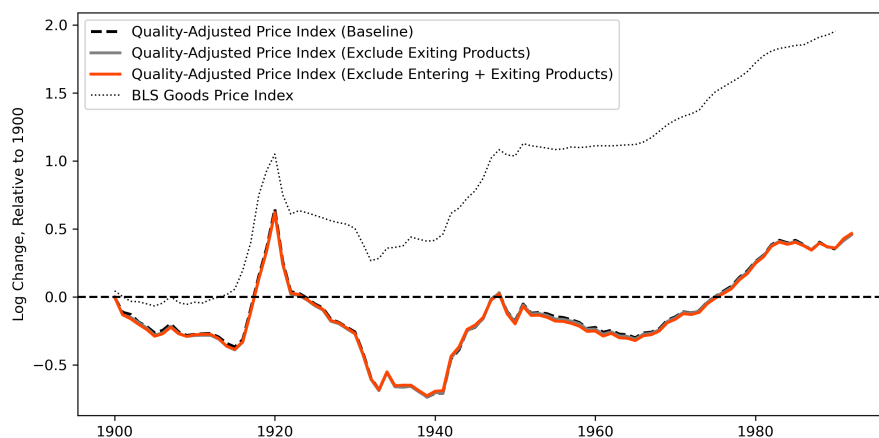
**Figure A.2:** Distribution of  $R^2$  For Hedonic Regressions, Log Differences



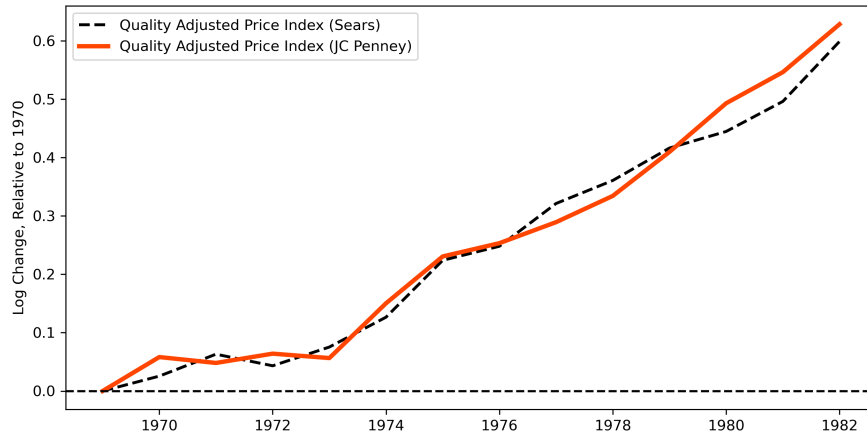
The figure shows kernel density estimates of in-sample and out-of-sample  $R^2$  from ridge regressions of annual matched-model log price changes on product characteristics, estimated separately by product category and year. Both distributions are restricted to regressions with positive out-of-sample  $R^2$ . The dashed vertical lines indicate the medians for this restricted sample, which are also reported in the legend.

**Figure A.3:** Sears Product Shares vs. CPI Weights

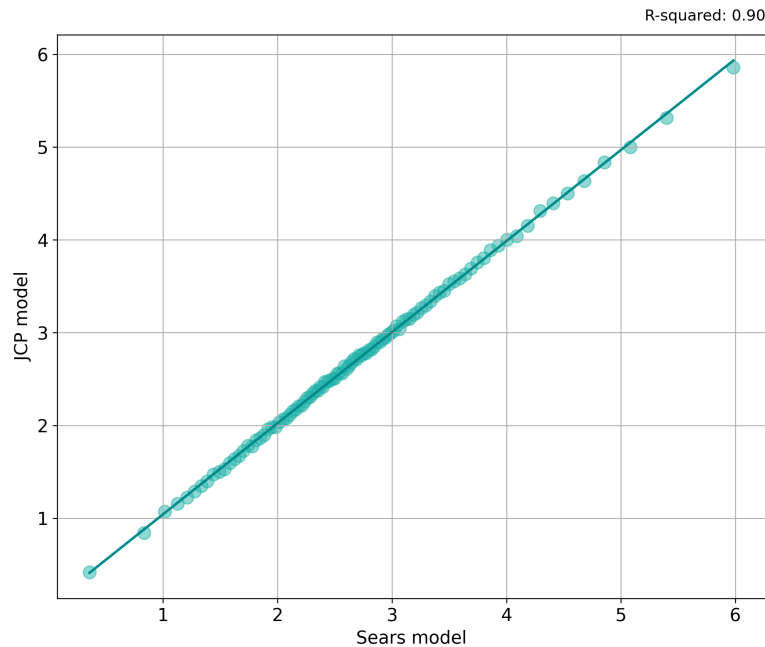
The figure is a bar plot comparing the shares of product listings across major BLS retail goods categories in the Sears catalogs with the corresponding nationally representative BLS expenditure weights. The Sears shares are based on product frequencies rather than quantities sold.

**Figure A.4:** Price Indexes, Restricting to Stable Products

The figure compares our baseline quality-adjusted Sears price index (dashed black line) with two indexes constructed after excluding products with no sufficiently similar counterpart in an adjacent year. The solid gray line excludes exiting products, whose closest substitute in the subsequent year has a text-embedding cosine similarity below 0.70. The solid red line excludes both exiting and entering products, with entry defined analogously using the preceding year. The dotted black line shows the historical BLS index for housefurnishings and apparel. The Sears indexes use the same BLS expenditure weighting procedure.

**Figure A.5:** Price Indexes, Sears vs. J.C. Penney

The figure presents quality-adjusted price indexes constructed from J.C. Penney catalogs (solid red line) and Sears catalogs (dashed black line) over 1969–1982. The Sears sample is restricted to Spring catalogs, and the comparison includes only product category–year cells observed in both retailers. The samples are matched at the category–year level, not at the level of identical individual products. Both indexes use the same BLS expenditure weighting procedure.

**Figure A.6:** Predicted Prices of J.C. Penney Products, Sears vs. J.C. Penney Hedonic Coefficients

The figure is a binscatter comparing predicted prices on the same product characteristics based on hedonic ridge regressions estimated separately using Sears and J.C. Penney catalogs for 1969–1982. We exclude the top and bottom 0.5% of outliers. The binscatter has 100 bins.

Figure A.7: Example of Appliances from 1910



(a) Refrigerator

**THE WORLD'S BEST WASHING MACHINE**

**THESE WORDS**, "The World's Best Washing Machine," tell the whole story. If we were to use half the words in this catalog, the dictionary and a dozen pages of this catalog we could not describe our High Speed Wizard Washer any better.

**WHY WE KNOW THIS WASHER IS THE BEST.** It is now generally known and admitted by all those who are interested in the manufacture of washing machines that the high speed rotary type of machine is far superior to any other type; that any machine of the high speed type will do far better work than the old rubbing machines, the wringer machines, the cylinder machines or any other type. We know that there would be an enormous demand for a high speed machine which would do the washing satisfactorily and which could be sold at a reasonable price. We satisfied ourselves that the high speed type of machine was the best, and we then ordered samples of nearly every high speed machine on the market. We put these machines in our laundry here at our plant and for days carefully tested each one. We compared their way running abilities; we compared the materials of which they were made, their appearance, and, in fact, every detail of their construction and operation. There was one machine which was far superior to all others in its ease of operation and the thoroughness with which it cleaned the clothes without injuring them. We went to the manufacturers of this machine and told them frankly that, in our opinion, they had the best washing machine on the market. We told them that we would contract to take an enormous number of these machines, provided they could save a very small profit on each machine, which would enable us to list it in our catalog at a price representing only a very small profit on the actual cost to manufacturers. When we had contracted to take now much money they could save by eliminating their traveling salesmen, their bad accounts and other expenses connected with selling machines through dealers, and told them how many machines we would take, they were that at the end of the year they would make more money by selling this enormous number of machines to us at a very small profit on each machine than they would if they sold only a comparatively few machines at a large profit on each machine through the jobbers and dealers. Consequently we contracted for the exclusive sale of this machine and we now offer it in this catalog as the best washing machine on the market, regardless of price.

**CONSTRUCTION**  
A HIGH GRADE WASHING MACHINE MUST COMBINE ALL THE FOLLOWING FEATURES TO INSURE PERFECT RESULTS AND LONG LIFE:

It is the action of water without rotating manner that will hold it under all circumstances. We use the best selected of the staves are made on automatic press. The tub is bound together by a galvanized wire hoops in center and of high grade close grain gray iron. It is from heavy unannealed patterns to High Speed Wizard Washer are of the same shape, accurately machined. This is the reason why the High Speed Wizard Washer is so superior to all other washers.

**WORKS**

is simple gearing on our High Speed Washer the long heavy casting by means of the heavy casting works on a station after inside of tub. When the handle turns back and forth inside of the tub, clothes through the hot ends and, in correct action, and bottom of tub. A pulley connects the handle wheel to a very large wheel on the inside of the tub. It will all jerky motion and when run it speed a slight push or pull on the handle is all that is needed to keep the machine. The fly wheel helps run the same manner that the balance of a run the coin shaker. The special set enables you to operate the High and Washer either sitting or standing. Only high speed water which can be either sitting or standing position and the only advantage we machine had it would pay you to get ours.

**USED FOR POWER.** The High and Washer can be operated by gasoline or by wheel with a 1-inch hole. If you would try wheel, suitable for a 5-inch hole.

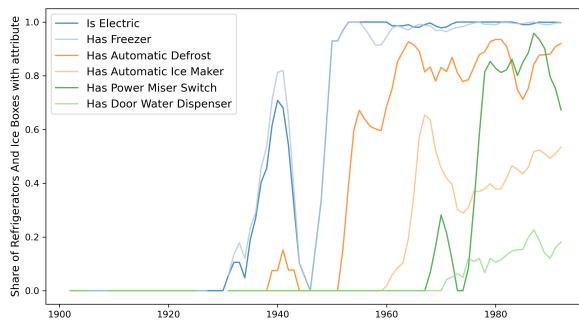
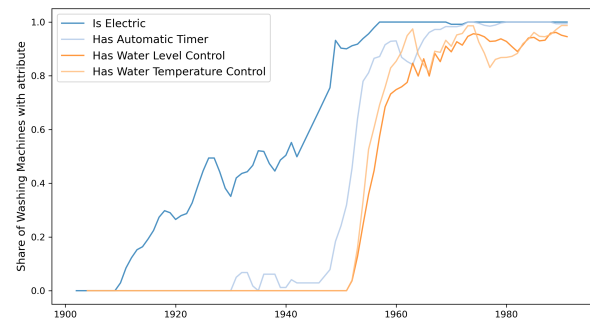
**TRIAL.** After you have used to every the machine and put it every statement which we have made true in every particular. If you do not find the machine to be the best washing machine you ever used or ever saw, and if you do not find that you have made a big saving, you may return it to us and get all your money back, including the freight charge. We want to keep the machine you will not be out one penny.

**\$7 15**

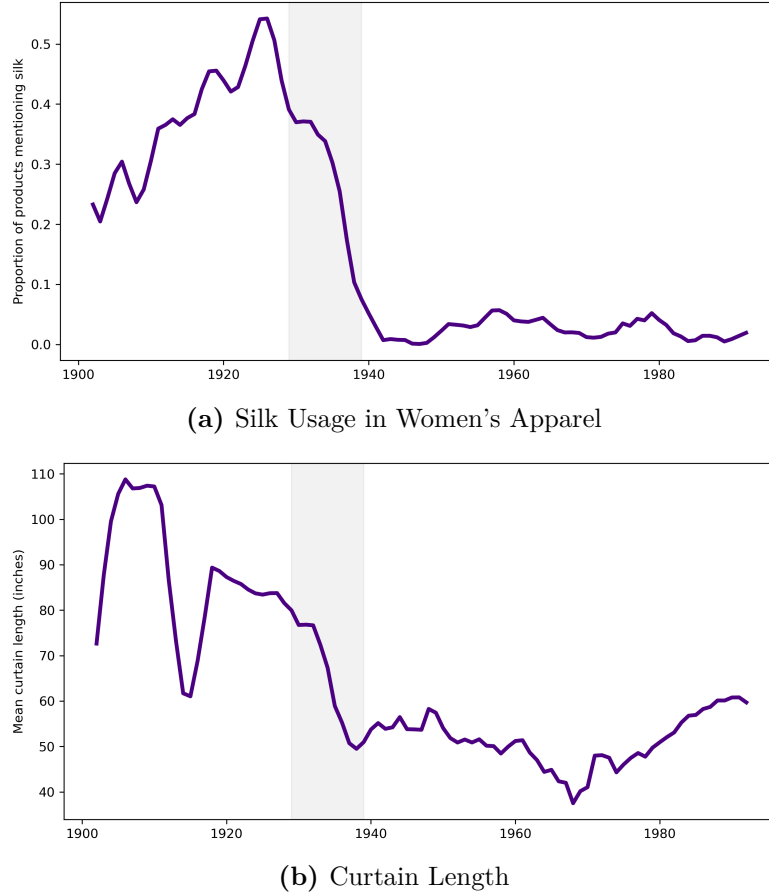
No. 1V7513

(b) Washing Machine

The figure presents sections from the 1910 Sears catalog featuring in Panel (a) a refrigerator and in Panel (b) a washing machine. Neither product was electric. The refrigerator used blocks of ice to cool the cabinet, while the washing machine was operated manually, as shown in the image.

**Figure A.8:** Diffusion of Electricity Into Appliances**(a)** Refrigerator**(b)** Washing Machine

The figure presents the diffusion of several technological advances in refrigerators and washing machines over time. Panel (a) shows the share of refrigerators listed in the Sears catalog that featured electricity, freezers, automatic defrosting, an ice maker, a power miser switch, and a door water dispenser. Panel (b) shows the share of washing machines that featured electricity, an automatic timer, water level control, and water temperature control. All lines are three-year moving averages.

**Figure A.9:** Quality Declines During the Great Depression

The figure shows two dimensions along which quality changed over the twentieth century. Panel (a) shows the proportion of women's apparel products (shirts, dresses, jackets, skirts, pajamas and sweaters) that used silk. Silk was a common material even at the start of the sample and its usage grew to more than 50% of women's apparel in the 1920s. Throughout this period, it was a luxury material and commanded a price premium of 0.8 log points on average. Silk usage collapsed during the Great Depression and remained low for the rest of the century. Panel (b) shows average curtain length in the Sears catalog. Before the Great Depression, curtains tended to be close to wall-length at around 80 inches (with the exception of WWI). During the Great Depression, average curtain length fell to approximately 50 inches, closer to the typical window size in the US. Both panels show 3-year moving averages.

## A.2 Digitizing Data Using LLMs

Our methodology requires extracting detailed structured information from more than one hundred thousand pages. Such a project would cost hundreds of thousands of dollars using conventional professional digitization services (Bäcker-Peral et al., 2025). For that reason, we use Gemini 3 Pro — the most powerful LLM for OCR tasks at the time of writing — to extract this information in a cost efficient manner. On average, digitization uses 3,000 input tokens and 9,000 output tokens per page, which at the current Gemini 3 API rate totals a cost of about \$0.06 per page — effectively two orders of magnitude lower than the cost of outsourcing to a digitization company.

For each page in the catalog, we use the following prompt to extract structured information from the text.

You are an expert data extraction agent specializing in historical retail catalogs.  
You are processing a page from a {company} catalog from {year}.

Inputs: Images of the catalog page (which you must analyze directly).

Your task is to analyze the images and to identify every unique product. Extract its key details, grouping all variants under a single parent product. Variants must have the same description and brand and must be the same type of product but can have minor differences.

Output Requirements: You must return only a valid JSON object. Do not include any introductory text, explanations, or apologies.

The JSON should have the following structure:

```
{
  "products": [
    {
      "name": "Product Name",
      "description": "General description",
      "short_description": "Summary",
      "hs_code": "123456",
      "is_accessory": false,
      "brand_name": "Brand",
      "versions": [
        {
          "name": "Version Name",
          "price": "$X.XX",
          "installments_price": "$X.XX",
          "downpayment_price": "$X.XX",
          "monthly_price": "$X.XX",
          "attributes": {
            "weight": "5kg"
          },
          "product_code": "SKU123"
        }
      ]
    }
  ]
}
```

Guidelines:

**products:** A list containing one object for each distinct product group found on the page. If no products are found, return an empty array: []. If the page is a glossary or index page that does not list specific products, return an empty array.

**products.name:** The main, overarching name for the product (e.g., "Craftsman Electric Drill", "Women's Day Dress").

**products.description:** The FULL description of the product, using the exact language from the catalog. Do not skip or abbreviate any parts of the description. Include slogans or parts of the description that are purely for advertisement purposes.

**products.short\_description:** A one sentence summary of the product in your own words

`products.hs_code`: Your best inference of the {hs\_year} 6-digit Harmonized System (HS) code for this product category. Do not classify it as an antique. Use your expert knowledge to carefully pick the code that best describes this product's purpose and specific features.

`products.is_accessory`: true or false

-true: If the item is an accessory (e.g., "replacement belt," "drill bits," "sewing machine needles") that is used with another primary product

-false: If the item is a standalone product (e.g., "drill," "dress," "refrigerator").

`products.brand_name`: The name of the brand of the product (e.g. "Kodak"). Use "" if unknown.

`products.versions`: A list containing all specific buyable variants of the parent product.

- Each variant must belong to the specific product category of the parent product and should all have the same HS code.

- Accessories for the product (like cases or replacement parts) must be listed as a different parent product.

- If a product has a table of sizes (e.g., 6ft, 8ft, 10ft) with different prices, each row is a separate Version.

- If a product has multiple colors but ONE price and ONE catalog number, treat it as a single Version with a "colors" attribute.

`products.versions.name`: The specific name or identifier for this version (e.g., "1/2-HP Model").

`products.versions.price`: The price for this specific version, extracted exactly as a string (e.g., "\$19.95"). If a cash and installments price is listed, use the cash price.

`products.versions.installments_price`: The price for this specific version if paid in installments. This is sometimes called the 'easy payment' price. If an installment price is not listed, use the empty string "".

`products.versions.downpayment_price`: If `installments_price` exists, the downpayment that must be paid at the time of purchase.

`products.versions.monthly_price`: If `installments_price` exists, the monthly payment that must be paid every month until repayment.

`products.versions.attributes`: A dictionary of key attributes of this specific version. Extract these directly from the image. If any of the following attributes are available, always include them:

- `shipping_weight`

- `warranty_length`

- `material`

- `origin_country`

- `num_units` (The number of identical copies of the product that come in the package. This will typically be 1.)

- `includes_accessories`: a comma-delineated list with any accessories that the product comes with

Include all attributes that are relevant to the buyer's decision, especially those which are not included in the product description but appear in the image or in a table.

products.versions.product\_code: The specific catalog number for this version

Other rules:

- Ensure all keys and string values in the JSON are enclosed in double quotes.
- When creating a string value that has a quotation inside the string, you must not use double quotes (") if the string is delimited with double quotes. Instead, convert the quotes inside the string to single quotes (').

CORRECT: "description": "A 3-button 'brain' for switching modes."

INCORRECT: "description": "A 3-button "brain" for switching modes."

- Return nothing but a JSON object. The output MUST be a valid JSON object.

We manually audited hundreds of pages and found low error rates in product extraction and price transcription.

### A.3 Details on the Hedonic Cost-of-Living Approximation

#### A.3.1 Welfare and the cost-of-living index

Consider a household  $h$  with utility function  $U^h(\mathbf{q})$ , where  $\mathbf{q}$  is its vector of consumption quantities. In period  $t$ , the household chooses its consumption bundle to maximize utility subject to its budget constraint:

$$\mathbf{q}_t^h \in \arg \max_{\mathbf{q} \geq 0} \left\{ U^h(\mathbf{q}) : \mathbf{p}_t \cdot \mathbf{q} \leq x_t^h \right\}, \quad (\text{A.1})$$

where  $\mathbf{p}_t$  is the vector of prices and  $x_t^h$  denotes household expenditure. Let  $u_t^h = U^h(\mathbf{q}_t^h)$  denote the utility attained by the household.

We can express welfare in expenditure units using the expenditure function,

$$e^h(\mathbf{p}, \bar{u}) \equiv \min_{\mathbf{q} \geq 0} \left\{ \mathbf{p} \cdot \mathbf{q} : U^h(\mathbf{q}) \geq \bar{u} \right\}. \quad (\text{A.2})$$

The expenditure function gives the minimum expenditure required to attain utility  $\bar{u}$  at prices  $\mathbf{p}$ . In particular,  $x_t^h = e^h(\mathbf{p}_t, u_t^h)$ .

We define the household-level cost-of-living index as the ratio of expenditures required to attain base-period utility under current and base-period prices:

$$COLI_t^h \equiv \frac{e^h(\mathbf{p}_t, u_{t-1}^h)}{e^h(\mathbf{p}_{t-1}, u_{t-1}^h)}. \quad (\text{A.3})$$

This measures the change in the minimum expenditure required for household  $h$  to maintain its base-period utility.

To aggregate across households, let  $\lambda_h$  denote the household's share of aggregate base-period expenditure,

$$\lambda_h \equiv \frac{x_{t-1}^h}{\sum_g x_{t-1}^g}. \quad (\text{A.4})$$

The expenditure-weighted aggregate cost-of-living index is then

$$\begin{aligned} COLI_t &\equiv \sum_h \lambda_h COLI_t^h \\ &= \frac{\sum_h e^h(\mathbf{p}_t, u_{t-1}^h)}{\sum_h e^h(\mathbf{p}_{t-1}, u_{t-1}^h)}. \end{aligned} \quad (\text{A.5})$$

The aggregate index measures the change in total expenditure required to maintain every household at its base-period utility level.

### A.3.2 The first-order Laspeyres approximation

The cost-of-living index in Equation (A.3) depends on household preferences, which are generally unobserved. We therefore consider an approximation that holds the base-period consumption bundle fixed.

Let  $\mathbf{q}_{t-1}^h$  denote household  $h$ 's expenditure-minimizing bundle at prices  $\mathbf{p}_{t-1}$  and utility  $u_{t-1}^h$ . By the definition of the expenditure function,

$$e^h(\mathbf{p}_{t-1}, u_{t-1}^h) = \mathbf{p}_{t-1} \cdot \mathbf{q}_{t-1}^h. \quad (\text{A.6})$$

By Shephard's lemma, the derivative of the expenditure function with respect to prices equals the vector of Hicksian quantities:

$$\nabla_{\mathbf{p}} e^h(\mathbf{p}_{t-1}, u_{t-1}^h) = \mathbf{q}_{t-1}^h. \quad (\text{A.7})$$

A first-order Taylor expansion of the expenditure function around base-period prices therefore gives

$$\begin{aligned} e^h(\mathbf{p}_t, u_{t-1}^h) &\approx e^h(\mathbf{p}_{t-1}, u_{t-1}^h) + \nabla_{\mathbf{p}} e^h(\mathbf{p}_{t-1}, u_{t-1}^h)'(\mathbf{p}_t - \mathbf{p}_{t-1}) \\ &= \mathbf{p}_{t-1} \cdot \mathbf{q}_{t-1}^h + \mathbf{q}_{t-1}^h \cdot (\mathbf{p}_t - \mathbf{p}_{t-1}) \\ &= \mathbf{p}_t \cdot \mathbf{q}_{t-1}^h, \end{aligned} \quad (\text{A.8})$$

where the second line substitutes Equation (A.6) and Shephard's lemma.

Substituting Equation (A.8) into the numerator of Equation (A.3), and using Equation (A.6) for its denominator, gives

$$COLI_t^h \approx \frac{\mathbf{p}_t \cdot \mathbf{q}_{t-1}^h}{\mathbf{p}_{t-1} \cdot \mathbf{q}_{t-1}^h} \equiv L_t^h, \quad (\text{A.9})$$

where  $L_t^h$  is the household-level Laspeyres price index. Since the household can always purchase its base-period bundle at current prices,  $COLI_t^h \leq L_t^h$ , provided that bundle remains available. The Laspeyres index is therefore an upper bound on the true cost-of-living index, and the preceding derivation shows that it is a first-order approximation.

Let  $\Omega_{t-1}$  denote the set of products available in the base period, and define household  $h$ 's base-period expenditure share on product  $i$  as

$$s_{i,t-1}^h \equiv \frac{p_{i,t-1} q_{i,t-1}^h}{\sum_{\ell \in \Omega_{t-1}} p_{\ell,t-1} q_{\ell,t-1}^h}. \quad (\text{A.10})$$

Using these shares, we can rewrite the Laspeyres index in Equation (A.9) as

$$\begin{aligned} L_t^h &= \frac{\sum_{i \in \Omega_{t-1}} p_{it} q_{i,t-1}^h}{\sum_{i \in \Omega_{t-1}} p_{i,t-1} q_{i,t-1}^h} \\ &= \sum_{i \in \Omega_{t-1}} \frac{p_{i,t-1} q_{i,t-1}^h}{\sum_{\ell \in \Omega_{t-1}} p_{\ell,t-1} q_{\ell,t-1}^h} \frac{p_{it}}{p_{i,t-1}} \\ &= \sum_{i \in \Omega_{t-1}} s_{i,t-1}^h \frac{p_{it}}{p_{i,t-1}}. \end{aligned} \quad (\text{A.11})$$

Aggregating across households using the expenditure weights defined in Equation (A.5), we obtain

$$\begin{aligned}
COLI_t &= \sum_h \lambda_h COLI_t^h \\
&\approx \sum_h \lambda_h L_t^h \\
&= \sum_h \lambda_h \sum_{i \in \Omega_{t-1}} s_{i,t-1}^h \frac{p_{it}}{p_{i,t-1}} \\
&= \sum_{i \in \Omega_{t-1}} \left( \sum_h \lambda_h s_{i,t-1}^h \right) \frac{p_{it}}{p_{i,t-1}}.
\end{aligned} \tag{A.12}$$

Defining the aggregate base-period expenditure share as

$$s_{i,t-1} \equiv \sum_h \lambda_h s_{i,t-1}^h = \frac{\sum_h p_{i,t-1} q_{i,t-1}^h}{\sum_h x_{t-1}^h}, \tag{A.13}$$

we obtain

$$COLI_t \approx \sum_{i \in \Omega_{t-1}} s_{i,t-1} \frac{p_{it}}{p_{i,t-1}}. \tag{A.14}$$

Thus, to a first order, the aggregate cost-of-living index equals the base-period expenditure-weighted average of product-level price relatives.

### A.3.3 Product turnover and hedonic imputation

The preceding derivation assumes that every product available in period  $t - 1$  can also be priced in period  $t$ . This assumption is problematic when products enter and exit the market. Let  $\Omega_t$  denote the set of products available in period  $t$ , and let  $\mathcal{M}_t = \Omega_t \cap \Omega_{t-1}$  denote the products observed in both periods. A conventional matched-model index measures price changes only for products in  $\mathcal{M}_t$ . If continuing products are not representative of all base-period products, their price changes will not recover the first-order approximation in Equation (A.14).

Suppose now that households have preferences over both the quantities and the characteristics of goods, and that firms face production costs that depend on these characteristics. As shown by Rosen (1974), a competitive equilibrium in such a market can be characterized by a hedonic price schedule that maps product characteristics into prices. Let  $j(i) \in \mathcal{J}$  denote the category of product  $i$ , and let  $X_{it} \in \mathcal{X}^{j(i)}$  denote its vector of characteristics. The equilibrium price of product  $i$  in period  $t$  can then be written as

$$p_{it} = P_{j(i),t}(X_{it}), \tag{A.15}$$

where  $P_{j,t}(\cdot)$  is the hedonic price schedule for category  $j$  in period  $t$ .

We can now write the household's expenditure problem allowing it to choose both quantities and characteristics. Let  $b^h = \{(q_i^h, X_i^h)\}_{i \in \Omega_{t-1}}$  denote a household consumption bundle, where  $q_i^h$  is the quantity purchased and  $X_i^h$  is its vector of characteristics.

Let  $U^h(b^h)$  denote utility from this bundle. Given the collection of hedonic price schedules  $\mathcal{P}_t = \{P_{j,t}\}_{j \in \mathcal{J}}$ , the cost of purchasing  $b^h$  is

$$C_t(b^h) \equiv \sum_{i \in \Omega_{t-1}} q_i^h P_{j(i),t}(X_i^h). \tag{A.16}$$

The household's expenditure function can therefore be written as

$$e^h(\mathcal{P}_t, \bar{u}) = \min_{b^h} \left\{ C_t(b^h) : U^h(b^h) \geq \bar{u} \right\}. \tag{A.17}$$

This is the expenditure problem in [Equation \(A.2\)](#), with prices determined by the characteristics of the goods purchased.

Let  $b_{t-1}^h = \{(q_{i,t-1}^h, X_{i,t-1})\}_{i \in \Omega_{t-1}}$  denote the household's optimal base-period bundle, and let  $\tilde{b}_t^h = \{(\tilde{q}_{i,t}^h, \tilde{X}_{i,t}^h)\}_{i \in \Omega_{t-1}}$  denote the expenditure-minimizing bundle under period- $t$  price schedules that attains the same utility. By construction,

$$U^h(b_{t-1}^h) = U^h(\tilde{b}_t^h) = u_{t-1}^h, \quad (\text{A.18})$$

and

$$e^h(\mathcal{P}_t, u_{t-1}^h) = C_t(\tilde{b}_t^h). \quad (\text{A.19})$$

Assume that the relevant characteristic bundles remain feasible in both periods, that utility and hedonic price schedules are sufficiently smooth, and that the relevant quantity and characteristic choices are locally interior. Let  $\mu_t^h$  denote the multiplier on the utility constraint in [Equation \(A.17\)](#). The first-order conditions for the expenditure-minimizing bundle  $\tilde{b}_t^h$  are

$$P_{j(i),t}(\tilde{X}_{i,t}^h) = \mu_t^h \frac{\partial U^h(\tilde{b}_t^h)}{\partial q_i}, \quad (\text{A.20})$$

$$\tilde{q}_{i,t}^h \nabla_X P_{j(i),t}(\tilde{X}_{i,t}^h) = \mu_t^h \nabla_{X_i} U^h(\tilde{b}_t^h). \quad (\text{A.21})$$

The first condition equates the marginal cost of quantity to its marginal utility, expressed in expenditure units. The second equates the marginal cost of improving a product's characteristics to the household's marginal valuation of those improvements.

Collecting the first-order conditions for quantities and characteristics, we can write

$$\nabla_b C_t(\tilde{b}_t^h) = \mu_t^h \nabla_b U^h(\tilde{b}_t^h). \quad (\text{A.22})$$

Because the two bundles attain the same utility, as established in [Equation \(A.18\)](#), a first-order Taylor expansion of utility around  $\tilde{b}_t^h$  gives

$$\begin{aligned} 0 &= U^h(b_{t-1}^h) - U^h(\tilde{b}_t^h) \\ &\approx \nabla_b U^h(\tilde{b}_t^h)'(b_{t-1}^h - \tilde{b}_t^h). \end{aligned} \quad (\text{A.23})$$

Multiplying by  $\mu_t^h$  and using the first-order conditions in [Equation \(A.22\)](#), we obtain

$$\nabla_b C_t(\tilde{b}_t^h)'(b_{t-1}^h - \tilde{b}_t^h) = \mu_t^h \nabla_b U^h(\tilde{b}_t^h)'(b_{t-1}^h - \tilde{b}_t^h) \approx 0. \quad (\text{A.24})$$

Next, a first-order Taylor expansion of the cost of the base-period bundle around the compensated period- $t$  bundle gives

$$\begin{aligned} C_t(b_{t-1}^h) &\approx C_t(\tilde{b}_t^h) + \nabla_b C_t(\tilde{b}_t^h)'(b_{t-1}^h - \tilde{b}_t^h) \\ &\approx C_t(\tilde{b}_t^h) \\ &= e^h(\mathcal{P}_t, u_{t-1}^h), \end{aligned} \quad (\text{A.25})$$

where the second line follows from [Equation \(A.24\)](#), and the final equality follows from [Equation \(A.19\)](#).

The same result as in the Laspeyres derivation therefore holds when households choose both quantities and characteristics: to a first order, the minimum expenditure required to attain base-period utility equals the cost of the base-period bundle under current prices.

Since the base-period bundle is optimal under  $\mathcal{P}_{t-1}$ , we also have

$$e^h(\mathcal{P}_{t-1}, u_{t-1}^h) = C_{t-1}(b_{t-1}^h). \quad (\text{A.26})$$

Substituting [Equations \(A.25\)](#) and [\(A.26\)](#) into the definition of the household-level cost-of-living index in [Equation \(A.3\)](#), we obtain

$$\begin{aligned} COLI_t^h &= \frac{e^h(\mathcal{P}_t, u_{t-1}^h)}{e^h(\mathcal{P}_{t-1}, u_{t-1}^h)} \\ &\approx \frac{C_t(b_{t-1}^h)}{C_{t-1}(b_{t-1}^h)} \\ &= \frac{\sum_{i \in \Omega_{t-1}} q_{i,t-1}^h P_{j(i),t}(X_{i,t-1})}{\sum_{i \in \Omega_{t-1}} q_{i,t-1}^h P_{j(i),t-1}(X_{i,t-1})}, \end{aligned} \quad (\text{A.27})$$

where the last line substitutes the definition of bundle cost from [Equation \(A.16\)](#).

Multiplying and dividing each term in the numerator by its base-period price gives

$$\begin{aligned} COLI_t^h &\approx \sum_{i \in \Omega_{t-1}} \frac{q_{i,t-1}^h P_{j(i),t-1}(X_{i,t-1})}{\sum_{\ell \in \Omega_{t-1}} q_{\ell,t-1}^h P_{j(\ell),t-1}(X_{\ell,t-1})} \frac{P_{j(i),t}(X_{i,t-1})}{P_{j(i),t-1}(X_{i,t-1})} \\ &= \sum_{i \in \Omega_{t-1}} s_{i,t-1}^h \frac{P_{j(i),t}(X_{i,t-1})}{P_{j(i),t-1}(X_{i,t-1})}, \end{aligned} \quad (\text{A.28})$$

where the second line uses the expenditure-share definition in [Equation \(A.10\)](#), together with  $p_{i,t-1} = P_{j(i),t-1}(X_{i,t-1})$  from [Equation \(A.15\)](#).

Aggregating across households as in [Equation \(A.5\)](#), and using the aggregate expenditure weights from [Equation \(A.13\)](#), yields

$$\begin{aligned} COLI_t &= \sum_h \lambda_h COLI_t^h \\ &\approx \sum_h \lambda_h \sum_{i \in \Omega_{t-1}} s_{i,t-1}^h \frac{P_{j(i),t}(X_{i,t-1})}{P_{j(i),t-1}(X_{i,t-1})} \\ &= \sum_{i \in \Omega_{t-1}} s_{i,t-1} \frac{P_{j(i),t}(X_{i,t-1})}{P_{j(i),t-1}(X_{i,t-1})}. \end{aligned} \quad (\text{A.29})$$

The price relative in the final expression evaluates both periods' price schedules at the same base-period characteristics. The denominator is observed for every base-period product, since  $P_{j(i),t-1}(X_{i,t-1}) = p_{i,t-1}$ . The numerator, however, may be unobserved when the product is no longer sold in period  $t$ .

For each base-period product, define its counterfactual current-period price as

$$p_{i,t}^* \equiv P_{j(i),t}(X_{i,t-1}). \quad (\text{A.30})$$

Substituting this definition and the observed base-period price into [Equation \(A.29\)](#) gives

$$COLI_t \approx \sum_{i \in \Omega_{t-1}} s_{i,t-1} \frac{p_{i,t}^*}{p_{i,t-1}}. \quad (\text{A.31})$$

To estimate  $p_{i,t}^*$ , let  $k(i)$  denote the most similar product of the same category  $j(i)$  observed in period  $t$ . Define the difference in characteristics between the base-period product and its closest substitute as

$$d_{it} \equiv X_{i,t-1} - X_{k(i),t}.$$

If the hedonic price schedule is smooth, a first-order Taylor expansion around the comparison product gives

$$\begin{aligned} p_{i,t}^* &= P_{j(i),t}(X_{i,t-1}) \\ &\approx P_{j(i),t}(X_{k(i),t}) + d'_{it} \nabla_X P_{j(i),t}(X_{k(i),t}) \\ &= p_{k(i),t} + d'_{it} \nabla_X P_{j(i),t}(X_{k(i),t}), \end{aligned} \quad (\text{A.32})$$

where the last line uses Equation (A.15) for the observed comparison product. The counterfactual price is therefore approximated by the observed price of the closest substitute, adjusted for the value of the differences in characteristics.

In our empirical specification, we assume that the hedonic price schedule is log-linear in characteristics, as in Equation (3.10):

$$P_{j,t}(X) = \exp \left( \delta_t^{j,t} + X' \beta^{j,t} \right). \quad (\text{A.33})$$

Under this functional form, substituting Equation (A.33) into the definition of the counterfactual price in Equation (A.30) gives

$$\begin{aligned} p_{i,t}^* &= \exp \left( \delta_t^{j,t} + X'_{i,t-1} \beta^{j,t} \right) \\ &= \exp \left( \delta_t^{j,t} + X'_{k(i),t} \beta^{j,t} \right) \exp \left( (X_{i,t-1} - X_{k(i),t})' \beta^{j,t} \right) \\ &= p_{k(i),t} \exp \left( (X_{i,t-1} - X_{k(i),t})' \beta^{j,t} \right). \end{aligned} \quad (\text{A.34})$$

Replacing  $\beta^{j,t}$  with the estimated coefficients  $\hat{\beta}^{j,t}$  from Equation (3.10), we obtain the empirical imputation

$$\hat{p}_{i,t} = p_{k(i),t} \exp \left( (X_{i,t-1} - X_{k(i),t})' \hat{\beta}^{j(i),t} \right). \quad (\text{A.35})$$

We can substitute these estimates of  $p_{i,t}^*$  into Equation (A.31) to construct the  $COLI_t$ .

#### A.4 Measuring Observed Quality Change

The hedonic price schedules derived in the preceding section can also be used to measure the economic value of changes in product characteristics. Consider a household  $h$  consuming product  $i$  with quantity  $\tilde{q}_{i,t}^h$  and characteristics  $\tilde{X}_{i,t}^h$ , as part of the compensated bundle  $\tilde{b}_t^h$  defined in Equation (A.17). Suppose that the product's characteristics improve by a vector  $\Delta X_i$ , while its quantity, purchase expenditure, and the household's other consumption remain initially fixed. We are interested in the household's willingness to pay for this improvement.

Let  $V_{i,t}^h$  denote the amount of expenditure the household would be willing to give up in exchange for the improvement, holding its utility constant. To express this compensation in dollars, suppose that the household also consumes a numeraire good  $c$  with price normalized to one. The household must be indifferent between receiving the quality improvement and giving up  $V_{i,t}^h$  units of the numeraire, and retaining its original bundle. A first-order Taylor expansion of this indifference condition around  $\tilde{b}_t^h$  gives

$$0 \approx \nabla_{X_i} U^h(\tilde{b}_t^h)' \Delta X_i - \frac{\partial U^h(\tilde{b}_t^h)}{\partial c} V_{i,t}^h. \quad (\text{A.36})$$

Rearranging, the household's willingness to pay is

$$V_{i,t}^h \approx \frac{\nabla_{X_i} U^h(\tilde{b}_t^h)' \Delta X_i}{\partial U^h(\tilde{b}_t^h) / \partial c}. \quad (\text{A.37})$$

We can express this willingness to pay in terms of the hedonic price schedule using the household's first-order conditions. Since the price of the numeraire is one, the first-order condition for numeraire consumption implies

$$1 = \mu_t^h \frac{\partial U^h(\tilde{b}_t^h)}{\partial c}, \quad (\text{A.38})$$

where  $\mu_t^h$  is the multiplier on the utility constraint in Equation (A.17). Substituting Equation (A.38) into Equation (A.37) gives

$$V_{i,t}^h \approx \mu_t^h \nabla_{X_i} U^h(\tilde{b}_t^h)' \Delta X_i. \quad (\text{A.39})$$

From the first-order condition for characteristics in Equation (A.21), we also know that

$$\mu_t^h \nabla_{X_i} U^h(\tilde{b}_t^h) = \tilde{q}_{i,t}^h \nabla_X P_{j(i),t}(\tilde{X}_{i,t}^h).$$

Substituting this expression into Equation (A.39), we obtain

$$V_{i,t}^h \approx \tilde{q}_{i,t}^h \nabla_X P_{j(i),t}(\tilde{X}_{i,t}^h)' \Delta X_i. \quad (\text{A.40})$$

Thus, the household's marginal willingness to pay for a change in product characteristics equals the corresponding change in the hedonic price schedule, multiplied by the quantity purchased. This is the equilibrium relationship established by Rosen (1974): at the household's optimal choice, the marginal price of a characteristic equals its marginal valuation in expenditure units.

We next express this valuation relative to expenditure on the product. Dividing Equation (A.40) by  $\tilde{q}_{i,t}^h P_{j(i),t}(\tilde{X}_{i,t}^h)$  gives

$$\begin{aligned} \frac{V_{i,t}^h}{\tilde{q}_{i,t}^h P_{j(i),t}(\tilde{X}_{i,t}^h)} &\approx \frac{\nabla_X P_{j(i),t}(\tilde{X}_{i,t}^h)' \Delta X_i}{P_{j(i),t}(\tilde{X}_{i,t}^h)} \\ &= \nabla_X \log P_{j(i),t}(\tilde{X}_{i,t}^h)' \Delta X_i. \end{aligned} \quad (\text{A.41})$$

The gradient of the log hedonic price schedule therefore measures the marginal willingness to pay for changes in characteristics, expressed relative to expenditure on the product.

Under the log-linear specification in Equation (A.33),

$$\log P_{j,t}(X) = \delta_t^{j,t} + X' \beta^{j,t},$$

the gradient of the log price schedule is simply

$$\nabla_X \log P_{j,t}(X) = \beta^{j,t}. \quad (\text{A.42})$$

Substituting Equation (A.42) into Equation (A.41) gives

$$\frac{V_{i,t}^h}{\tilde{q}_{i,t}^h P_{j(i),t}(\tilde{X}_{i,t}^h)} \approx (\Delta X_i)' \beta^{j(i),t}. \quad (\text{A.43})$$

Hence, the hedonic coefficients provide a first-order approximation to the household's willingness to pay for a small quality improvement, as a fraction of expenditure on the product.

We can aggregate this valuation across products within a category. Let  $x_{j,t}^h$  denote household  $h$ 's expenditure on category  $j$  at its compensated period- $t$  bundle,

$$x_{j,t}^h \equiv \sum_{i:j(i)=j} \tilde{q}_{i,t}^h P_{j(i),t}(\tilde{X}_{i,t}^h), \quad (\text{A.44})$$

and define the expenditure share of product  $i$  within category  $j$  as

$$\omega_{i,t}^h \equiv \frac{\tilde{q}_{i,t}^h P_{j,t}(\tilde{X}_{i,t}^h)}{x_{j,t}^h}. \quad (\text{A.45})$$

Let  $V_{j,t}^h$  denote the household's willingness to pay for a collection of small quality improvements  $\{\Delta X_i\}$  within category  $j$ . Since the first-order valuations are additive across products, summing Equation (A.43) gives

$$\begin{aligned} \frac{V_{j,t}^h}{x_{j,t}^h} &\approx \sum_{i:j(i)=j} \omega_{i,t}^h (\Delta X_i)' \beta^{j,t} \\ &= \left( \sum_{i:j(i)=j} \omega_{i,t}^h \Delta X_i \right)' \beta^{j,t}. \end{aligned} \quad (\text{A.46})$$

As in our price index, we assume equal expenditure weights across products within each category, and use nationally representative expenditure weights to aggregate across categories. We also assume that changes in the characteristics of products offered in the Sears catalog are representative of changes in the characteristics of products consumed.

Let  $\Omega_{j,t}$  denote the set of products in category  $j$  observed in the catalog in year  $t$ , and define their average characteristics as

$$\tilde{X}_{j,t} \equiv \frac{1}{|\Omega_{j,t}|} \sum_{i \in \Omega_{j,t}} X_{it}. \quad (\text{A.47})$$

Under these assumptions, the expenditure-weighted change in characteristics in Equation (A.46) is approximated by the difference between the average characteristics of products offered in each year:

$$\sum_{i:j(i)=j} \omega_{i,t}^h \Delta X_i \approx \tilde{X}_{j,t} - \tilde{X}_{j,t-1}. \quad (\text{A.48})$$

Substituting Equation (A.48) into Equation (A.46) gives

$$\frac{V_{j,t}^h}{x_{j,t}^h} \approx (\tilde{X}_{j,t} - \tilde{X}_{j,t-1})' \beta^{j,t}. \quad (\text{A.49})$$

In practice, some characteristics cannot be determined from every product description. For each characteristic  $\ell$  in category  $j$ , let  $\pi_{j\ell,t}$  denote the share of products for which the characteristic is observed, and let  $\bar{X}_{j\ell,t}^{obs}$  denote its mean among those products. We assume that missing characteristic values in both periods can be imputed using the base-period observed mean,  $\bar{X}_{j\ell,t-1}^{obs}$ . Under this assumption, the change in the average characteristic is, to a first order, equal to the change among observed products weighted by their base-period reporting share:

$$\tilde{X}_{j\ell,t} - \tilde{X}_{j\ell,t-1} \approx \pi_{j\ell,t-1} (\bar{X}_{j\ell,t}^{obs} - \bar{X}_{j\ell,t-1}^{obs}). \quad (\text{A.50})$$

22

Applying Equation (A.50) to each characteristic and substituting into Equation (A.49), we obtain

$$\frac{V_{j,t}^h}{x_{j,t}^h} \approx \sum_{\ell=1}^{K_j} \pi_{j\ell,t-1} (\bar{X}_{j\ell,t}^{obs} - \bar{X}_{j\ell,t-1}^{obs}) \beta_{\ell}^{j,t}, \quad (\text{A.51})$$

<sup>22</sup>To see this, let  $\tilde{X}_{j\ell,t}$  denote the mean after imputing missing values with  $\bar{X}_{j\ell,t-1}^{obs}$ . Then  $\tilde{X}_{j\ell,t} = \pi_{j\ell,t} \bar{X}_{j\ell,t}^{obs} + (1 - \pi_{j\ell,t}) \bar{X}_{j\ell,t-1}^{obs}$ , while  $\tilde{X}_{j\ell,t-1} = \bar{X}_{j\ell,t-1}^{obs}$ . Consequently,  $\tilde{X}_{j\ell,t} - \tilde{X}_{j\ell,t-1} = \pi_{j\ell,t} (\bar{X}_{j\ell,t}^{obs} - \bar{X}_{j\ell,t-1}^{obs})$ . Replacing  $\pi_{j\ell,t}$  with  $\pi_{j\ell,t-1}$  introduces only the product  $(\pi_{j\ell,t} - \pi_{j\ell,t-1})(\bar{X}_{j\ell,t}^{obs} - \bar{X}_{j\ell,t-1}^{obs})$ , which is second order for small annual changes. Thus, the approximation does not require reporting shares to remain constant.

where  $K_j$  is the number of characteristics in category  $j$ .

Replacing the population coefficients with the estimates  $\hat{\beta}^{j,t}$  from Equation (3.10), we define our measure of observed quality change as

$$q_{j,t} = \sum_{\ell=1}^{K_j} \pi_{j\ell,t-1} \left( \bar{X}_{j\ell,t}^{obs} - \bar{X}_{j\ell,t-1}^{obs} \right) \hat{\beta}_{\ell}^{j,t}. \quad (\text{A.52})$$

When all characteristics are observed,  $\pi_{j\ell,t-1} = 1$  for every  $\ell$ , and this reduces to

$$q_{j,t} = (\tilde{X}_{j,t} - \tilde{X}_{j,t-1})' \hat{\beta}^{j,t}, \quad (\text{A.53})$$

as in Equation (5.1). In our implementation, we use three-year moving averages of the observed characteristic means to smooth discontinuities in the catalog data before computing annual changes.

Finally, we aggregate across categories using nationally representative expenditure shares. Let  $s_{j,t-1}$  denote category  $j$ 's share of national expenditure in the base period. Since Equation (A.51) expresses the welfare value of quality change relative to category expenditure, multiplying by  $s_{j,t-1}$  converts it into a contribution relative to aggregate expenditure. Summing these contributions gives

$$Q_t = \sum_{j \in \mathcal{J}} s_{j,t-1} q_{j,t}. \quad (\text{A.54})$$

Substituting Equation (A.52), we obtain

$$Q_t = \sum_{j \in \mathcal{J}} s_{j,t-1} \sum_{\ell=1}^{K_j} \pi_{j\ell,t-1} \left( \bar{X}_{j\ell,t}^{obs} - \bar{X}_{j\ell,t-1}^{obs} \right) \hat{\beta}_{\ell}^{j,t}. \quad (\text{A.55})$$

Under the maintained assumptions,  $Q_t$  provides a first-order approximation to the gross willingness to pay for improvements in observed product characteristics, expressed as a fraction of aggregate expenditure. The measure values changes in characteristics using a common hedonic price schedule and therefore isolates the economic value of quality improvements from changes in the overall level of prices.

## A.5 Extracting Structured Features from the Data

Our analysis in Section 3.1 requires detailed measures of product characteristics that are not directly available as structured variables in the catalog data. We use large language models (LLMs) to extract these characteristics from product descriptions. Our procedure consists of two steps. First, we construct a set of relevant attributes for each product category. Second, we use these attributes to extract the characteristics of individual products.

To construct the attribute lists, we provide an LLM with a sample of product descriptions from each category and ask it to identify characteristics that are important for distinguishing products. We then expand these lists incrementally over time. We first collect a set of characteristics that are relevant for the full sample period. Then, for each decade, we provide the model with a sample of products from that period, together with the attributes already identified, and ask it to propose additional characteristics that are not captured by the existing list. We limit the number of attributes that can be added in each iteration and instruct the model to avoid redundant or overly specific characteristics. This procedure allows us to capture technological changes over time, including characteristics that may have been important only during particular historical periods, while maintaining a consistent set of attributes across the entire sample.

We restrict the attributes to objectively verifiable characteristics that can be represented as either binary or numeric variables. For example, materials and product features are represented using binary indicators, while dimensions, capacity, and weight are represented using numeric variables. We instruct the

model to exclude brand names, colors, subjective descriptions, and attributes that duplicate information already captured by other variables. The resulting attribute lists provide a common set of characteristics with which to compare products within each category, including products sold many decades apart.

Once we have constructed these attribute lists, we query an LLM again to extract the corresponding values for individual products. For each product, we provide the model with its name, version name, description, and year, together with the complete list of attributes and their definitions.

For each attribute, the model first uses information explicitly stated in the product description or unambiguously implied by it. When this information is insufficient, we allow the model to make informed inferences using other product characteristics and its knowledge of historical products. We instruct the model to interpret characteristics according to their historical context rather than imposing contemporary product standards. When the available information does not permit a reasonable inference, the attribute is recorded as missing.

The model returns numeric values in standardized units and binary values indicating the presence or absence of each characteristic. We use a structured output format to ensure that the extracted values conform to the predefined attribute schema.