The Age Bias of Innovation

Kai Arvai[1] Antonin Bergeaud[2] Katja Mann[3] Rentian Zhu[4]

Preliminary Draft
March 25, 2026

Abstract

How does innovation respond to long-run shifts in aggregate demand? This paper introduces a new dataset that uses modern language-model-based methods to map the universe of U.S. patents into detailed consumption categories from 1971–2020. Leveraging profound changes induced by the demographic transition, we show that innovation has progressively reoriented toward goods consumed by older households. This shift in innovation activity generated price dynamics that disproportionately benefit older consumers, while raising the relative price of goods consumed by the young. We develop a model of directed innovation in which demographic aging endogenously redirects R&D effort and gives rise to increasing intergenerational consumption inequality.

Keywords: Innovation, AI, Demography, Consumption, Inflation
JEL Classification: E21, E22, J31, O33, O41

[1]European Central Bank. E-mail: kai.arvai@ecb.europa.eu
[2]HEC Paris and CEPR. E-mail: antonin.bergeaud@gmail.com
[3]Copenhagen Business School. E-mail: kma.eco@cbs.dk
[4]Copenhagen Business School. E-mail: rz.eco@cbs.dk

We thank Julia Schmidt for helpful discussions and suggestions. We furthermore thank Simon Schmitz for excellent research assistance. The views in this paper are of the authors and do not represent the views of the ECB or the Eurosystem.

1 Introduction

Between 1970 and 2020, the number of U.S. patents related to healthcare technologies grew by over 150 percent in share, while patents in clothing, furniture, and tobacco declined by 40 to 65 percent. Over the same period, the baby boomer generation aged from early adulthood into retirement, shifting the center of gravity of consumer demand toward the goods and services consumed later in life. Are these two trends connected? More generally, does the direction of technological innovation respond to changes in the composition of aggregate demand? A body of work has established that market size shapes innovation in specific sectors: Acemoglu and Linn (2004) show that demographic shifts redirect pharmaceutical patenting, Finkelstein (2004) documents similar effects in vaccines, and Clemens (2013) finds that Medicare spurred medical equipment innovation. These papers establish the market size channel convincingly in pharmaceuticals. But does the same logic extends across the full range of goods and services that households consume? If it does, then demographic change could reshape the entire landscape of innovation, with consequences for relative prices and welfare that sector-level analyses cannot fully capture.

In this paper, we assemble a new dataset that matches the universe of product patents granted between 1971-2020 to a large and representative set of consumption items and analyze how the nature of innovation evolves over time when aggregate demand is changing. Leveraging the profound changes induced by demographic aging in the U.S., we show that the overall direction of technological change has shifted: an increasing share of patenting activity is directed toward technologies that cater to the needs of an older population. We document which consumption and patent categories are driving this trend. We also show how this shift in the direction of technological progress has led to relative price movements that are more beneficial for the old, while causing higher relative prices for the young.

We make four main contributions. First, we develop a three-stage pipeline that matches patent abstracts to consumer product categories using transformer-based embeddings and large language models, validated against human annotations. The resulting dataset contains over 63 million patent-product links spanning nearly 6 million patents and 251 consumption categories. This provides, to our knowledge, the first comprehensive mapping between patented technologies and the goods and services that households actually consume.

Based on this dataset, we document a set of new stylized facts. There are vast differences in the applicability in the universe of patents across product categories, ranging from very narrow patents to patents that are widely applicable. Furthermore,

there has been a change in the patent shares of product categories over time. While in the 1970s, patents in the areas of transportation and home maintenance were most frequent, the focus had shifted towards electronic patents in the 2010s. The share of patents falling on healthcare, entertainment and financial products also grew over the sample period. In contrast, patents in tobacco, clothing and apparel and furniture saw a relative decline.

Second, the shift in patenting activity has consequences for relative prices: By linking our dataset to breakthrough innovation measures of Kelly et al. (2021), we show that patenting advances in a product area lead to a decline in that product's relative price. In local projections, relative price responses –while small and insignificant in the years immediately following the break-through innovation– become significant and negative after seven years. They reach a peak nine years after the invention, where a one standard deviation innovation shock leads to inflation being 0.8 percentage points lower than what it otherwise would have been. This accumulates to a substantial decline in the relative price of a product that is exposed to breakthrough innovation. We also document a hump-shaped increase in general patenting activity after such a shock. This is consistent with the innovation-competition channel documented by Jaravel (2019).

Third, by merging the resulting patent-product links with household-level consumption data from the Consumer Expenditure Survey, we construct, for each age group and year, a measure of "innovation exposure". This is defined as the patent intensity of the consumption basket. We show that the patent intensity differs systematically across age groups, increasing markedly between ages 20 and 60, and thereafter declining slightly. The age gradient of innovation has become steeper over time, indicating that patenting has become more tilted towards older age groups.

We show that underlying the steeper age gradient of patenting is the aging of the baby boomer cohorts. While in prime age in the early years of our sample, the baby boomers are close to retirement at the end of our sample. In a regression set-up controlling for the age-year-cohort problem, we show that both age and cohort effects are strongly significant, and that the cohort effect is strongest for households born in the large groups of the 1950s and early 1960s. This provides evidence that the size of a cohort matters for innovation, mirroring the finding by Acemoglu and Linn (2004). In support of this hypothesis, we show that cohort size emerges as a significant explanatory factor for patent intensity after controlling for age and year fixed effects.

Fourth, we propose a mechanism that rationalizes these findings and clarifies why the redirection of innovation toward older consumers is a recent phenomenon, why it generates intergenerational inequality, and why the market equilibrium is inefficient. The key insight is that the demographic transition interacts with fixed life-cycle con-

sumption patterns to produce a time-varying reallocation of R&D effort. Different age groups consume systematically different baskets, and these differences become most pronounced after age 50, when spending tilts sharply toward healthcare, financial services, and utilities, and away from education and housing. When a large cohort such as the baby boomers reaches this part of the life cycle, the effective market size of "old-consumer" sectors grows substantially. Profit-maximizing firms respond by directing innovation toward these categories. The resulting quality improvements and cost reductions benefit the large older cohort disproportionately, while sectors consumed by the young receive comparatively less innovation and experience higher relative price growth. This mechanism generates rising intergenerational consumption inequality endogenously, with the demographic transition as the sole exogenous driver. In Section 5, we develop this argument formally and discuss the sources of inefficiency that arise from market-directed innovation.

Literature Review Our paper connects to several strands of the literature.

First the literature on market size and directed innovation. The idea that the size of a market shapes the direction of technological change has a long tradition, from the Hicksian theory of induced innovation to Schmookler (1966), who first documented that patenting responds to demand conditions. Acemoglu (2002a) formalizes this insight in a general equilibrium framework, showing that an increase in the relative supply of a factor can redirect innovation toward technologies that complement it. The most direct empirical evidence comes from pharmaceuticals: Acemoglu and Linn (2004) exploit exogenous variation in U.S. demographic structure to show that increases in potential market size lead to more drug patenting. In a different setting, Aghion et al. (2024) show that increases in export market size raise domestic innovation, though only for the most productive firms.

Second, the literature on the innovation and inequality nexus. A large literature has focused primarily on the income distribution. On the supply side, Acemoglu (1998) and Acemoglu (2002b) argue that skill-biased technical change, itself a response to the growing supply of skilled labor, is a central driver of wage inequality. Autor et al. (2003) provide supporting evidence from labor markets. More recently, Acemoglu and Restrepo (2018) study the implications of automation technologies for the labor share and wage structure. On the firm side, Aghion et al. (2019) show that innovation contributes to top income inequality through entrepreneurial rents, while Autor et al. (2020) link the rise of superstar firms to technological concentration. Our paper shifts the focus from the income distribution to the age distribution, and from the labor market to the product market: we study how innovation affects inequality not through wages but through the

differential quality and price of goods consumed by different groups. This consumption-side perspective complements the production-side view that dominates the existing literature.

Third, a recent literature on product innovation, prices, and welfare. The channel through which innovation affects consumers most directly is through prices and product quality. Jaravel (2019) establishes the causal chain from product innovation to increased competition to lower relative prices, showing that innovation targeted at high-income consumers reduces the prices they face and thereby widens real consumption inequality. Argente et al. (2020) link patents to product descriptions in Nielsen scanner data and confirm that patenting firms are responsible for a large share of new product introductions. On the measurement side, a body of work documents that aggregate price indices mask substantial heterogeneity in the inflation rates faced by different groups: Hobijn and Lagakos (2009) show that household-specific inflation rates vary systematically with demographics and income; Kaplan and Schulhofer-Wohl (2017) highlight the importance of heterogeneous consumption baskets for understanding real inequality; and Jaravel and Lashkari (2024) demonstrate that standard welfare measures substantially misstate welfare growth across the income distribution. Furthermore, Arvai and Mann (2026) show how differential exposure to digital technology in the consumption basket creates price dynamics that are more advantageous for rich households. We extend the innovation-to-prices-to-inequality chain to the age dimension and show that the direction of innovation itself is endogenous to demographic structure.

Fourth, a growing literature examines how population aging affects the rate and direction of innovation. On the aggregate rate, Bloom et al. (2020) show that maintaining constant growth requires exponentially growing research inputs, implying that aging economies with stagnating R&D labor forces face declining growth. Aksoy et al. (2019) and Maestas et al. (2023) confirm empirically that aging reduces innovation and productivity growth. Jones (2009) provides the theoretical foundations for why demographic change constrains idea production. On the composition of innovation, Acemoglu et al. (2022) find that younger inventors are disproportionately responsible for radical breakthroughs, suggesting that an aging workforce may independently shift innovation toward incremental improvements. Acemoglu and Restrepo (2022) formalize the link between aging and the adoption of automation technologies. On the demand side, Bornstein (2025) shows that older households exhibit significantly greater brand loyalty and lower switching rates, creating a harvesting motive for incumbents that discourages new firm entry and accounts for approximately 20 to 30 percent of the observed decline in U.S. firm entry rates. This consumer inertia channel interacts with ours: even though older consumers represent a large market, its inertial nature may favor incremental quality improvements by incumbents over radical product innovation by

entrants, connecting to the broader literature on declining business dynamism (Akcigit and Ates, 2021; Aghion et al., 2023). Our paper adds a demand-pull channel to this literature: we show that aging reshapes not just who innovates or how productive they are, but what gets innovated, through the composition of consumer demand.

Finally, we contribute methodologically to the growing economics literature that exploits the textual content of patents. Kelly et al. (2021) use textual similarity between patents to identify breakthrough innovations over the long run. Bergeaud et al. (2017) classify patents based on semantic content. Mann and Püttmann (2023) identify automation innovations by training algorithms on patent texts. Kogan et al. (2023) link patents to worker-level data using text analysis. Bloom et al. (2021) combine patent texts with job postings and earnings calls to trace the diffusion of new technologies across firms and workers. Existing crosswalks between patent classifications and industry or product codes, such as those based on CPC codes or input-output tables (Lybbert and Zolas, 2014), are too coarse to capture which consumer goods benefit from a given technology. Our three-stage pipeline, combining domain-specific transformer embeddings with large language model reasoning and human validation, provides a finer-grained mapping from patents to the consumption basket. This complements Argente et al. (2020), who link patents to product-level Nielsen scanner data, and extends the scope to the universe of consumer expenditure categories.

The remainder of the paper is organized as follows. Section 2 describes the patent, consumption, and price data. Section 3 presents our patent classification pipeline and links it to consumption and price data, showing that relative prices decline when a prouduct is exposed to an innovation shock. Section 4 documents stylized facts on the direction of innovation and its age gradient, and tests whether patenting responds to cohort size. Section 5 develops the economic mechanism linking demographic change to directed innovation and intergenerational inequality. Section 6 concludes.

2 Data sources

2.1 Patent data

We use the universe of utility patents granted between 1971-2020 in the United States. From the PTO PatentsView database, we extract for each patent the title and abstract along with patent metadata including publication numbers, application and grant dates, and Cooperative Patent Classification (CPC) codes. We use the patent abstract as the basis for our patent-product classification. By PTO guidelines, the abstract is required to summarize the technical field, the solution, and the principal use(s) of the invention, and

hence contains the information most relevant for identifying which consumer products a patent could benefit. Using abstracts rather than full patent texts is standard in the economics literature on patent text analysis (e.g. Kelly et al., 2021; Bergeaud et al., 2017) and is computationally tractable at the scale of nearly 6 million patents.

2.2 Consumption data

To measure consumption of households, we use data from the Consumer Expenditure Survey (CEX). The CEX is the largest consumption survey in the U.S., carried out at the household level and encompassing several thousand households per wave. Next to providing detailed expenditure data, it also contains a host of demographic and income information of the households. We use the interview survey data of the public-use microdata, which covers the full consumption basket of households from 1984-2023. While consumption expenditures are available at detailed expenditure categories from 1990, between 1984-89, only summary expenditure measures are available. We follow Jaravel and Lashkari (2024) and interpolate between summary categories by assuming that the allocation of consumption within categories in these years is the same as in 1990.

We treat each survey wave as an independent observation, i.e. not summing up over the different survey rounds in which each household participates. We drop households with non-positive total household income as well as households with an age below 20 or above 80 due to the small number of observations in these age groups. The household age is the average age of household head and spouse (if present).[1] We calculate annual expenditures at the level of calendar years following the procedure suggested by the Bureau of Labor Statistics.

2.3 Price data

The price series are constructed using data from the BEA's National Income and Product Accounts (NIPA). We use NIPA Tables on personal consumption expenditures by type of product. Tables 2.4.3U, 2.4.4U, 2.5.4U refer to consumed quantities, prices and personal consumption expenditures, respectively. To match these with the consumption categories of the CEX, we use a crosswalk from NIPA PCE lines to UCC codes provided by the BLS and create Törnqvist indices for more aggregate UCC categories.

[1] The results are similar when using only the household head's age, where we follow the convention to define the man as the household head in a mixed-gender couple.

3 Methodology

3.1 Product categories

We use the product classification of the Consumer Expenditure Survey, the Universal Classification Codes (UCC). The CEX has about 600 different UCCs in each year, but for some categories the classification changes quite substantially across years. We therefore summarize these into 251 distinct and time-consistent UCCs (see Appendix B.2). The product categories are further divided into 15 summary categories. We define our own summary categories for this purpose, rather than relying on the 32 aggregate product categories from the CEX summary tables. Besides being on a more aggregate level, our categories group expenditures based on purposes, rather than industry. For example, we group insurance with the respective expenditure category that is being insured, and proceed in the same way with other financial charges. Rentals and repairs are merged with the product category they refer to. Our summary categories are: (1) Food & Beverages, (2) Housing & Shelter, incl. home insurance and mortgage payments, (3) Home Maintenance and Repairs, incl. home services, (4) Utilities, (5) Financial Services & Financial Charges, incl. insurances other than health, property or vehicle, (6) Household Furnishings, incl. dinnerware, (7) Clothing & Apparel (8) Healthcare, incl. personal care and health insurance, (9) Education and Childcare, (10) Transportation, incl. vehicle financial charges and insurance, (11) Entertainment & Recreation, incl. pets and travel and vacation, (12) Tobacco Products, (13) Household appliances, (14) Electronic equipment, (15) Other. A detailed list of UCCs is provided in Table B.1.

To extract a set of words that describe each UCC, we use the Information Booklet, which the BLS provides to survey respondents to assist them in remembering purchases, identifying eligible expenses, and ensuring accurate reporting. The Information Booklet contains lists and examples of items. In addition, we use ChatGPT to add further descriptions to these product categories. We ask the AI to first describe the consumption category and to give a brief overview of the production process[2]. For example, the consumption category "sofas" is described as "a furniture commonly used in living rooms, lounges, offices etc, that potentially multiple users can sit on". It then explains the production process, that entails designing, material collection, cutting, sewing, assembling and final finish.

[2]The exact prompt is: "Describe the consumption category '{word} and explain the production process."

3.2 Patent classification

Linking patents to consumption categories poses a dual challenge: the space of potential pairs is enormous (over 1.5 billion combinations across 6 million patents and 251 UCC categories), yet matching requires nuanced semantic judgment that simple keyword rules cannot provide. We address this with a three-stage pipeline that combines the scalability of vector search with the reasoning capabilities of large language models, validated against human annotation.

The first stage drastically reduces the search space while retaining semantically relevant candidates. We embed both patent abstracts and UCC category descriptions using PatentSBERTa (Bekamiri et al., 2021), a transformer model fine-tuned specifically on patent text. This choice is deliberate: general-purpose sentence encoders trained on web or news corpora perform poorly on the specialized, dense technical language of patent abstracts, whereas PatentSBERTa captures domain-specific terminology and inventive concepts. Each patent abstract p_i and category description c_j is mapped to a 768-dimensional vector, and we rank candidate pairs by cosine similarity,

$$S_{ij} = \frac{\mathbf{e}_i^p \cdot \mathbf{e}_j^c}{\|\mathbf{e}_i^p\| \, \|\mathbf{e}_j^c\|}, \tag{1}$$

using FAISS to perform approximate nearest-neighbor search efficiently at scale. For each patent we retain the 50 highest-scoring categories, reducing the problem from over 2 billion candidate pairs to roughly 300 million while preserving semantically plausible matches for downstream review.

Vector similarity alone, however, is insufficient for determining whether a patent technology would *meaningfully benefit* a consumption category: cosine distance captures surface-level semantic overlap but cannot reason about functional relevance or downstream applicability. The second stage therefore passes each pre-screened pair to GPT-4o-mini for binary classification. Each prompt presents the model with the full patent abstract, the 50 candidate categories from Stage 1, and explicit criteria distinguishing meaningful technological benefits from superficial thematic overlap. The model returns a binary decision for each candidate category, indicating whether the technology could plausibly improve products or services in that category. This distinction is the core methodological contribution of the classification stage: prior keyword-based and co-occurrence approaches ask whether a patent *relates* to a category, whereas we ask whether it could *benefit* one, a fundamentally different and more economically meaningful criterion. We select GPT-4o-mini over larger variants to optimize the precision-cost tradeoff at this scale, and process patents in batches of 1,000 via the OpenAI Batch API.

The third stage assesses classification quality through systematic human annotation of 735 randomly chosen patents. Research assistants are presented with each patent abstract alongside the AI-generated category assignments and asked to judge, for each assigned category, whether they agree that the technology could meaningfully benefit it. This design reflects how the pipeline will be used in practice: rather than asking annotators to construct classifications from scratch, we assess whether the AI output meets a reasonable standard of human judgment. Annotators follow detailed guidelines emphasizing direct applicability rather than thematic adjacency, and may override any AI decision they consider incorrect.

We evaluate agreement using Jaccard similarity, which measures set overlap between AI and human classifications at the patent level, as well as precision, recall, and F1 scores. The AI classifications achieve a mean Jaccard similarity of 72.7 percent and a median of 84.6 percent, with 73.6 percent of patents reaching strong agreement (Jaccard ≥ 0.60). Patent-level average precision is 81.3 percent, average recall is 79.8 percent, and average F1 is 78.5 percent. The near-equal precision and recall confirm that the system neither systematically over-assigns nor under-assigns categories relative to human judgment. At the category level, micro-averaged F1 reaches 85.2 percent, reflecting strong performance on commonly occurring categories; macro-averaged F1 is 68.7 percent, indicating room for improvement on niche categories with sparse training signal.

The pipeline produces 63.5 million patent-category links spanning over 6 million patents and 251 UCC categories. The validation results provide confidence that these links capture genuine pathways through which patented technologies diffuse into consumer goods and services, enabling the first systematic, large-scale analysis of innovation-to-consumption linkages at the category level.

3.3 Merging patent and consumption data

We merge patent and consumption data based on the UCC and the patent grant year. We calculate three different patent measures at the level of UCCs: a raw count of granted patents per year, a citations-weighted patent count that accounts for the importance of the patents, and a novelty measure based on the number of other patents that are cited in the patent. Concretely, we use the number of citations received in the first five years after the patent grant date as a measure of the patent's importance, and the number of backward citations that the patent uses as a measure of its novelty, using data by Kelly et al. (2021).

3.4 Stylized Facts

The overall number of patents in our final dataset has increased dramatically from 51,195 in 1971 to 299,843 in 2020, closely mirroring the general increase in patenting.[3] Out of the 5,913,835 patents in the final sample, 5,526,426 are matched to at least one UCC, and the number of matched patents tracks the total number of patents over time (Figure A.1). The share of matched patents increases from 0.91 in the 1970s to 0.95 in 2020 (Figure A.2). As shown in Table 1, for patents with at least one matched UCC, the average number of matches is 6.36, whereas the average number of matched supercategories is 2.49. There is large variation across patents though.

Table 1: UCC and supercat matches per patent

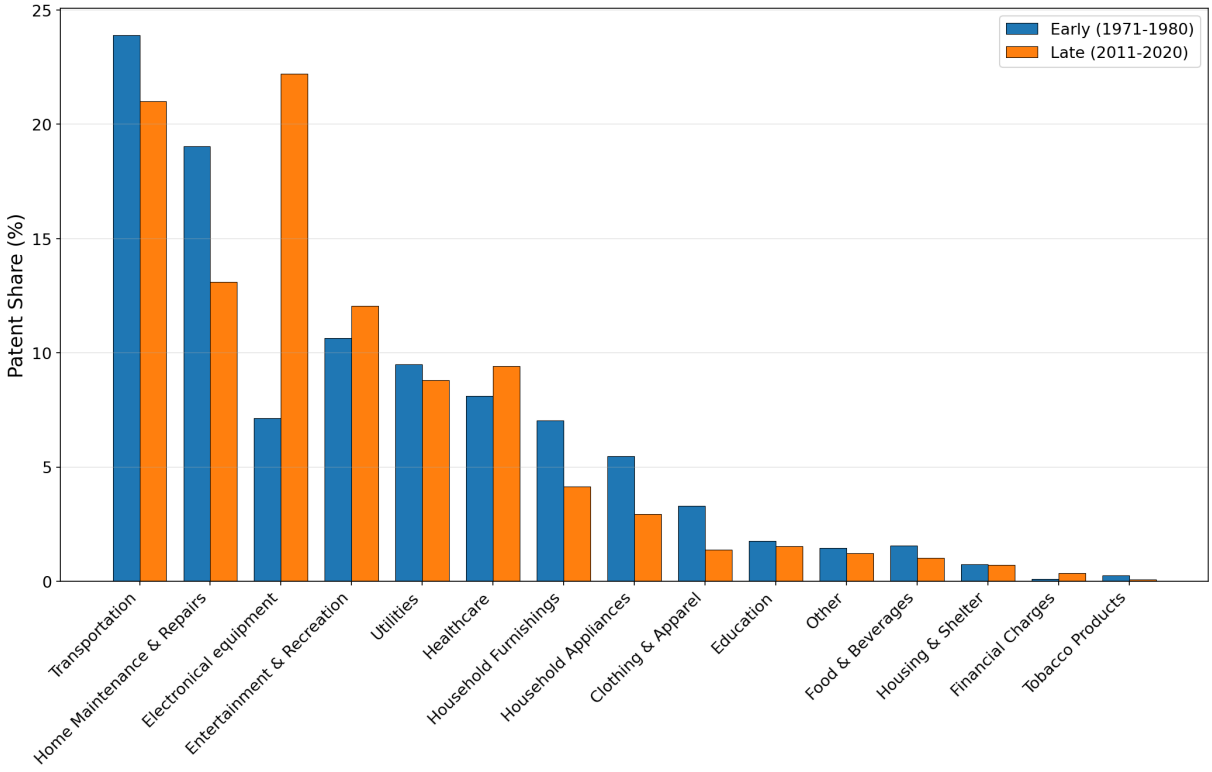
	mean	min	p25	p50	p75	max
UCC	6.36	1	4	6	8	45
supercategory	2.49	1	2	2	3	13

Over the sample period, the average number of UCC matches increased from 6.02 to 6.57 (Figure A.3). This means that not just are there more patents, the patents also have become more broadly applicable across final consumption categories. The increase is mostly concentrated between 1990 and 2008, which is related to the fact that in the 1990s, software became patentable. Software patents are often broadly applicable. The patent assigned the largest number of supercategories (13) and UCCs (45), i.e. almost universally applicable, is `Patent 9288053 - Schema Signing`, a patent on digital signature technology, granted in 2016 to Microsoft Technology Licensing LLC. Table A.2 lists all patents assigned to 40 or more UCCs. All of these 12 patents refer to digital inventions, covering areas such as e-commerce, electronic signal transmission, online payment systems or data management tools. All are granted in the 2000s and held by large corporations, most of them U.S.-based. This highlights the key importance of digital technology for new consumer product inventions, a point also made by Arvai and Mann (2026) in the context of consumption inequality.

Figure 1 shows which product categories are responsible for most patents. Blue bars show the initial period 1971-1980, whereas orange bars focuses on the more recent period between 2011 and 2020. As mentioned above, the most striking development has been an increase in patents on electronic equipment, whose share increased by 15 percentage points (209.7%). Other categories with increases are healthcare (17.6%), entertainment

[3]The total number of utility patents granted were 78,317 in 1971 and 352,049 in 2020. Detailed numbers across time can be found here.

Figure 1: Patent share in major categorizes over time

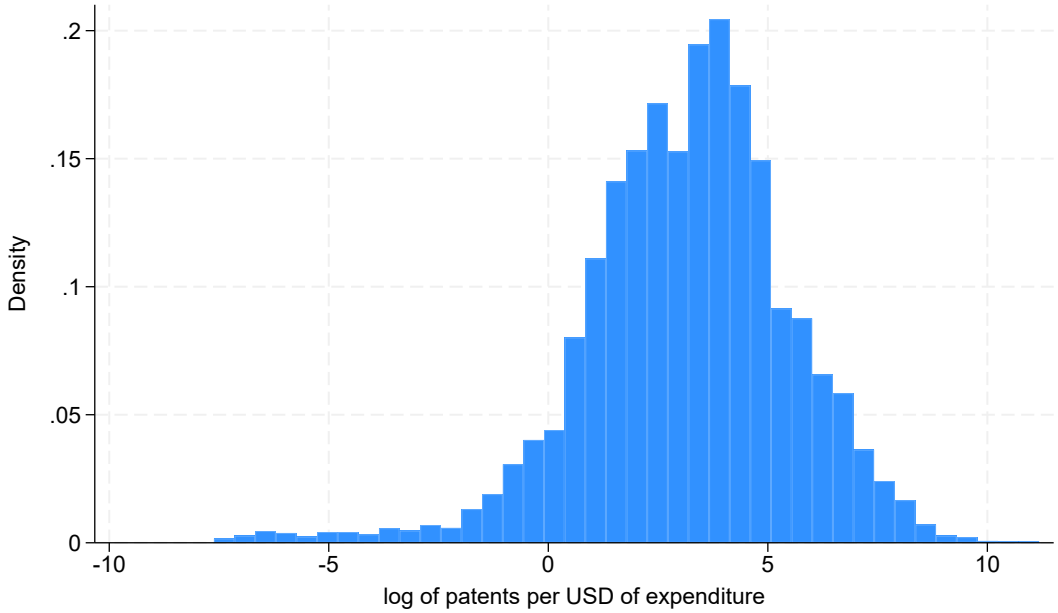


and recreation (13.2%), and financial charges and financial services (360.8%). The latter has the lowest absolute patent numbers in the early period, though. All other categories have lost in importance. The largest drop in patent shares happened in tobacco products (-65.7%), clothing and apparel (-58.5%), household appliances (-46.4%) and household furnishings (-41.2%).

Within electronics, the UCCs with the largest increase are computer hardware and software, video games, television, and other sound and video equipment. In healthcare, medical equipment, medical care services outside of hospitals, and services by medical professionals other than physicians increased the most.

Figure 2 shows the (log of the) number of patents in each UCC divided by the US Dollar amount spent on that UCC by an average consumer across all years in our sample. Patents are approximately log-normally distributed across expenditure, implying a large right tail. This is a first evidence that patent activity is indeed concentrated on larger markets in terms of USD expenditures.

Figure 2: Patents per USD spent



4 Empirical analysis

4.1 Effects of Innovation on Prices

This section documents a robust negative relationship between patenting activity and the inflation rate of a consumption category. We estimate impulse responses using the local projection method of Jordà (2005) to trace the evolution of UCC-level inflation, following an innovation shock.

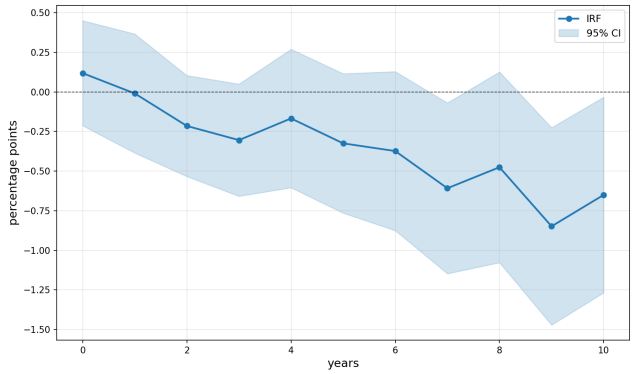
Our analysis is conducted at the level of consumption categories indexed by c and years indexed by t. For each UCC-year pair we compute the inflation rate $\pi_{c,t}$. We focus on patents classified by Kelly et al. (2021) as breakthrough innovations. They classify patents as breakthroughs based on textual similarity to previous and subsequent work: If a patent is distinct from previous work but related to subsequent innovations, it is identified as a break-through. These are patents that capture unexpected technological advances rather than incremental innovation driven by market size. As such, they provide a more plausibly exogenous source of variation in technological progress[4].

We estimate the following sequence of local projections:

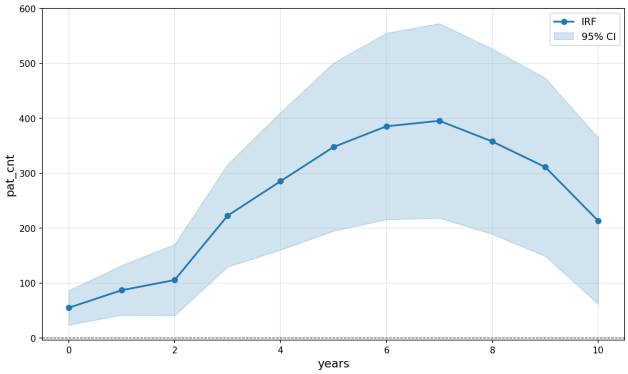
[4]Figure A.5 provides a summary on the distribution of these break-through on major consumption categories

Figure 3: Impulse response to an innovation shock

(a) Inflation



(b) Patenting



Note: This graph displays the impulse response function of inflation to a one standard deviation increase of break-through patenting activity, using regression (2). The graph in panel (b) replaces inflation $\pi_{c,t}$ with the pure patent count in the regression. Standard errors are clustered.

$$\pi_{c,t+h} = \beta_h \log(1 + Breakthrough_{c,t}) + \rho\pi_{c,t-1} + \alpha_c + \gamma_t + \varepsilon_{c,t+h}, \tag{2}$$

for horizons $h = 0, 1, \ldots, 10$. Here, $Breakthrough_{c,t}$ denotes the number of break-through innovations associated with UCC c in year t. The coefficients β_h trace the impulse response of category-level inflation h years after an innovation shock. The specification includes UCC fixed effects α_c and year fixed effects γ_t. UCC fixed effects absorb time-invariant differences in inflation dynamics across consumption categories, while also controlling for different patenting levels across UCCs, for example because innovations in some consumption categories are generally more patentable than in others. Year fixed effects capture aggregate macroeconomic shocks that affect all categories simultaneously, such as economy-wide inflation or business cycle fluctuations. Including these fixed effects ensures that identification relies on differential movements in prices across categories within a given year. To account for persistence in category-level price dynamics, we additionally control for lagged inflation $\pi_{c,t-1}$. Standard errors are clustered at the category level to allow for serial correlation within consumption categories over time.

Figure 3 (a) presents the resulting impulse response function. The estimated response of inflation is close to zero in the year of the shock. Over the subsequent years, however, the response becomes increasingly negative, reaching its largest magnitude approximately nine years after the innovation shock with -0.8 percentage points. This implies that after a one standard deviation shock in the break-through innovation measure for a certain category, inflation in year 9 is 0.8 percentage points lower than it would otherwise have been. This accumulates to a substantial relative price decline of that

category over 10 years due to a break-through innovation.

This dynamic pattern is consistent with a mechanism in which breakthrough innovations trigger a sequence of follow-on innovations and technological improvements that reduce production costs over time. We confirm this hypothesis with another local projection of breakthrough innovation on total patent numbers in a UCC. The IRF reveals a hump-shaped increase in patenting activity after break-through patents have laid the groundwork, see Figure 3 (b). Rather than affecting prices immediately, innovation appears to exert downward pressure on category-level prices with a delay, reflecting the time required for technological advances to diffuse into production processes and market outcomes.

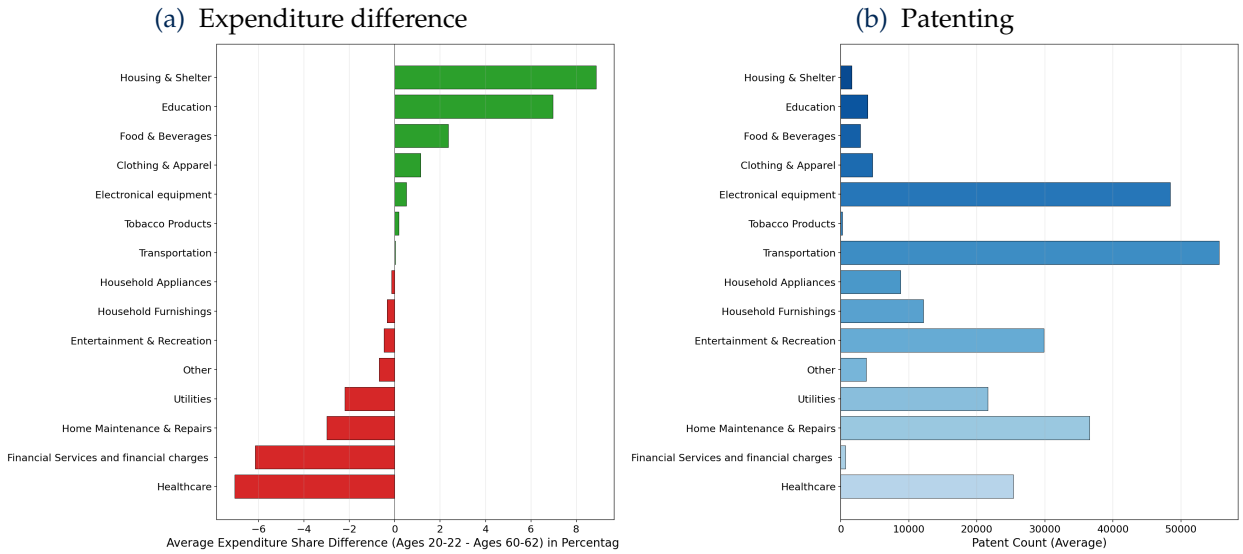
A purely descriptive correlation of average decadal inflation rates and the average level of patenting activity also confirms that a negative relationship holds consistently across decades and becomes even more pronounced when aggregating to broader super-categories, as indicated in Figure A.10 . These findings suggest that patenting intensity captures innovative activity that exerts downward pressure on relative prices, consistent with models in which technological progress reduces production costs and intensifies competitive forces.

4.2 Age patterns of consumption

We begin by showing that there is an age component to consumption, i.e. that the consumption baskets of households vary across age. This point has been made e.g. by Aguiar and Hurst (2013).

Figure 4 (a) shows average percentage point differences in expenditure shares between households of age 20-22 and age 60-62 over the entire time horizon. Negative values refer to categories with a higher consumption share for older households, whereas positive values indicate a stronger weight in the consumption basket of younger households. More relevant for the young than for the old are in particular Housing & Shelter (+8.5pp) and Education (+7pp), and also Food & Beverages (+2.3pp). Those categories are very low in terms of patenting activity, as indicated in Figure 4 (b). Note that the relevance of Housing & Shelter is partly due to the lower homeownership rate of 20-year olds compared to 60-year olds. Those who own a house and have paid off their mortgage will have very little expenditure falling in the Housing and Shelter category. Instead, their housing-related expenses will rather show up in the Home Maintenance & Repairs category (-3pp). Older households spend also relatively more on Healthcare (-7pp), Financial Services & Financial Charges (-6pp) and Utilities (-2pp). Note that healthcare and utilities are in particular patent intensive. Interestingly, while electronics

Figure 4: Age differences in expenditure and patenting



Note: Panel (a) shows the expenditure share difference by super category for households aged 20-22 and 60-62 for the aveage between 1984-2020 in percentage points, while panel (b) shows the average patenting count per year between 1984-2020.

and transportation are as discussed the most innovative consumption categories, the expenditure difference between young and old in these categories are small. Figure A.4 shows more detailed life-cycle consumption patterns for selected product categories. The most pronounced life-cycle pattern emerges in Healthcare and Education, which is hardly surprising.

4.3 Patent intensity across age groups

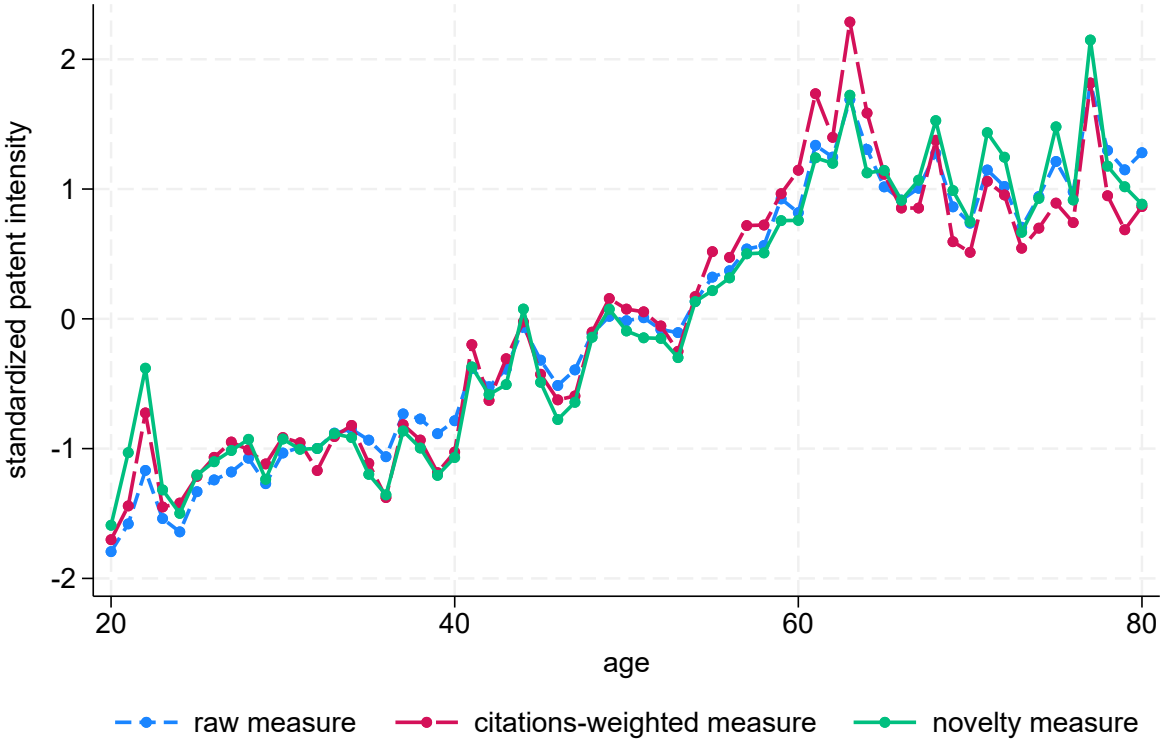
We compute the patent intensity in the consumption basket of individual household as a weighted mean of the UCC-level patent measures, where we use the individual household's expenditure shares of each UCC as weights. Then we create a year- and age/cohort-specific panel dataset, where expenditures are weighted averages of all households of the respective age and year, using the survey weights. The patent intensity is then calculated by using the age group/cohort's expenditure shares to weigh UCC-specific patent measures.

We compute the patent-intensity of an age group by multiplying the patent measure of a consumption category with the average expenditure share of that item, and summing those up across all consumption of a group. When plotting a standardized patent intensity across age without including any controls or addressing the correlation with cohorts, there emerges a pronounced life-cycle pattern of patenting. Figure 5 shows

numbers for three different patent measures, the raw measure, the citations-weighted and the novelty measure. We are showing standardized measures as they have vastly different units. There is a clear positive age gradient of patent intensities up until about age 60. Afterwards, the pattern flattens out. The figure resembles life-cycle consumption profiles (Gourinchas and Parker, 2002; Fernández-Villaverde and Krueger, 2007; Aguiar and Hurst, 2013), although these patterns usually find a steeper decline after retirement, mirroring the decline in income. Nevertheless, this figure suggest a connection between expenditure and patenting. Figure A.7 shows the same graph but instead of using patent intensities, it shows the residual of a regression of patent intensities on year fixed effects and a set of controls: mean household income, mean education, and the average urban share. The age pattern is still clearly visible, and even becomes slightly more pronounced. So this pattern is not driven by changes in income over the life-cycle. Figure A.8 shows that the age pattern is also visible if only using patents held by public companies as classified by Dyevre and Seager (2023), which presumably are to a larger extent guided by profit-maximizing motives. In contrast, there is no age pattern when focusing only on breakthrough patents. The innovations underlying these patents should to a lesser extent be invented out of commercial concerns.

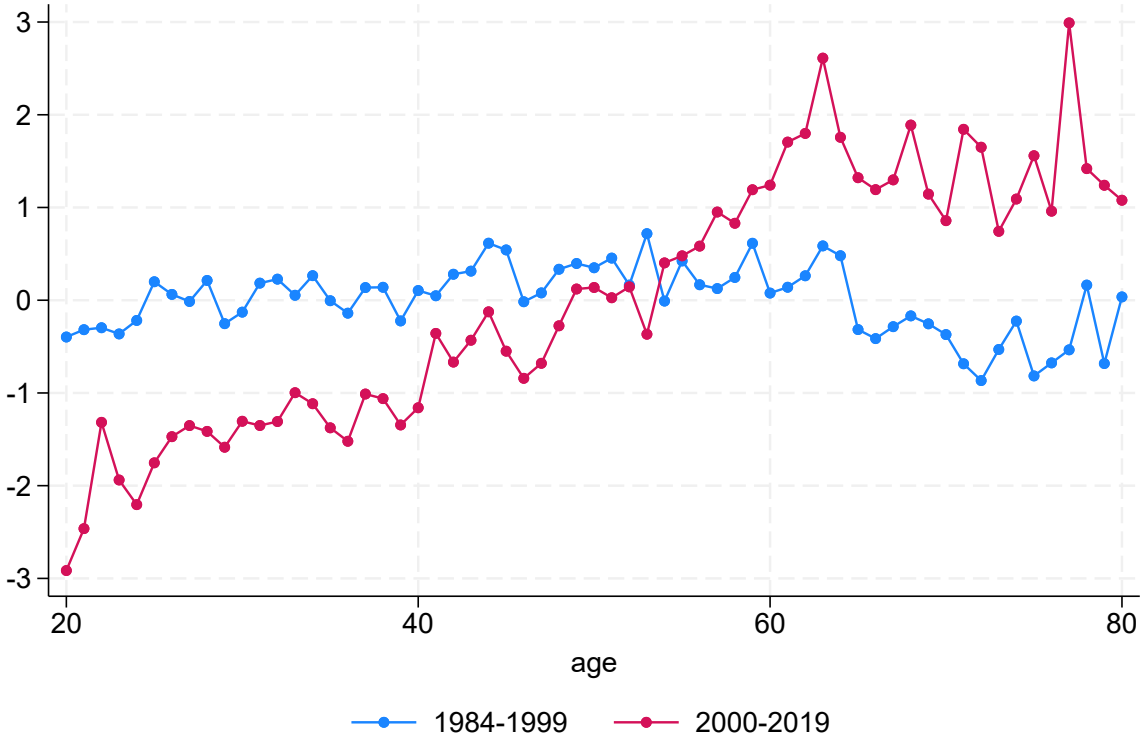
Figure 6 shows that the pronounced slope in age is only visible for half of the sample. In 1984-99, patent intensity is only slightly increasing until about age 50, and after that slightly declining. In 2000-2020, there is a pronounced positive slope until about age 60, and a flattening out or small decline afterwards. We do not observe anything comparable when studying patenting across income groups. This suggests that something has changed in the patenting strategies of inventors. One particular feature is the change in the size of age groups across time. Figure A.6 shows that the population mass moves from peaking at around age 20 in 1980 to peaking at around age 60 in 2020. This is due to the aging of the baby boomer cohort (those born between 1946 and 1964). This cohort constitutes a large market, and is therefore of interest to inventors.

Figure 5: Patents across age



Note: Standardized patent intensities, averages across 1984-2019.

Figure 6: Patents across age - subperiods



Note: This panel shows the residualized age component for two sub periods. The regression is

4.4 Patents across cohorts

To formally test the connection between patenting and cohort size, we carry out regressions at the household-level and at the age/cohort-level.

Consumption patterns contain age-, cohort- and time-specific elements. We are facing the issue well known in the life-cycle literature (e.g. Ameriks and Zeldes, 2002, Fagereng et al., 2017) that, without additional restrictions, it is not possible to identify these effects separately. This is due to the fact that these age, year and cohort are perfectly correlated. Potential solution strategies include imposing parametric restriction, or replacing one set of dummies with variables capturing the causal mechanism underlying that effect. In the following, we will apply both strategies. Using the first solution, we follow Deaton and Paxson (1994) and impose the restriction that time effects sum to zero once the variables have been detrended. Alternatively, we include a linear time trend. Using the second solution, we replace cohort dummies by the size of the cohort.

We test whether cohort size predicts patent intensity in the following regression

$$\text{patint}_{cat} = \text{cohort size}_{ct} + \alpha_a + \alpha_t + X_{cat} + \epsilon_{cat} \tag{3}$$

where c denotes cohort, a age and t year. *patint* is the patent intensity measure. In the first specification, *cohort size* is a simple measure of the number of people in a specific cohort alive at year t. Alternatively, we multiply this number with age-specific mean income and mean expenditure (calculated as simple averages of real income or real expenditures across all years) to approximate purchasing power. We include age and year fixed effects α_a and α_t as well as a set of controls X, which includes mean income, mean education and the share of households living in an urban area.

Table 2 shows the regression results using different patent measures: the raw patent measure, an importance-weighted measure (weighting patents by citations in the five following years) and the novelty-weighted measure, measuring the number of patents cited by this patent. The coefficient on *cohort size* is always positive and significant.

In Figure 7, we address the cohort-age-time problem differently. We estimate

$$\text{patint}_{cat} = \alpha_a + \alpha_t + X_{cat} + \epsilon_{cat} \tag{4}$$

by imposing that time effects sum to zero once the variables have been detrended. We include the same controls as before. Then, we plot the estimated standardized residual $\hat{\epsilon}_{cat}$ of this estimation against cohorts. Figure A.9 is the equivalent for including a linear time trend. For comparison, we show the cohort size in million of citizens.

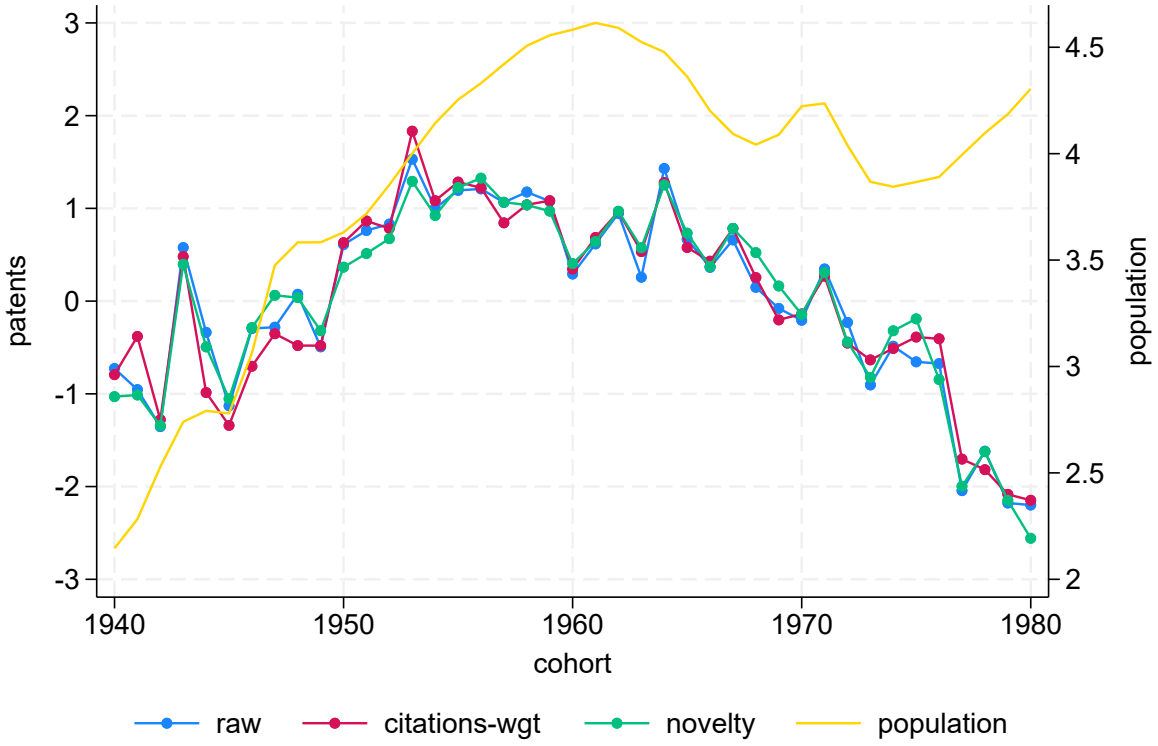
We find strong evidence that patenting activity indeed tracks the large baby boomer

Table 2: Patents and cohort size

Variables	(1) raw	(2) raw	(3) cit-wgt	(4) cit-wgt	(5) novelty	(6) novelty
cohort size	60.11***	47.00***	901.4***	733.0***	594.4***	503.6***
	(7.375)	(7.052)	(90.51)	(89.44)	(70.11)	(67.61)
Constant	266.3***	-682.2***	-3,236***	-13,160***	-1,241***	-10,396***
	(38.79)	(128.9)	(476.0)	(1,635)	(368.7)	(1,236)
Observations	2,196	2,196	2,196	2,196	2,196	2,196
R-squared	0.991	0.992	0.995	0.995	0.994	0.995
Controls	no	yes	no	yes	no	yes
Year FE	yes	yes	yes	yes	yes	yes
Age FE	yes	yes	yes	yes	yes	yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 7: Patent intensity across cohorts - demeaned



cohort, focusing the innovation effort of the economy on the largest and most wealthy group.

4.5 Age Cohorts and Inflation Inequality

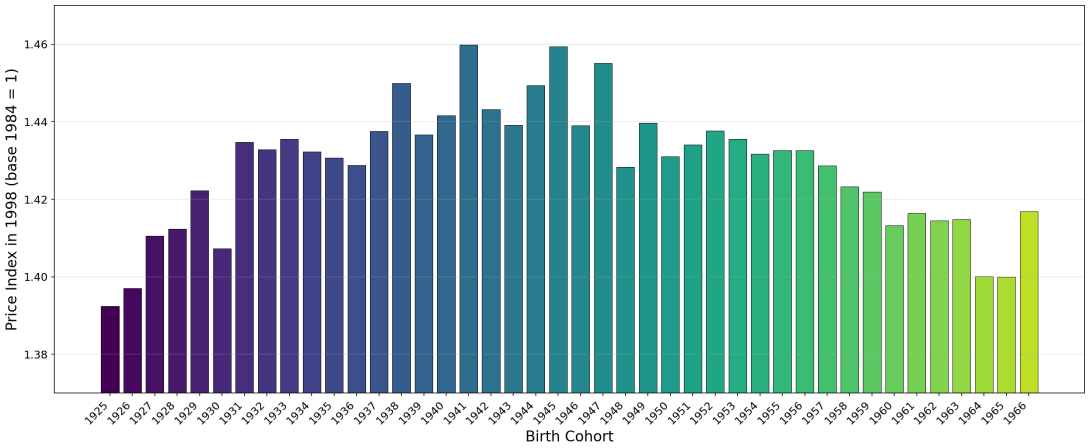
This section connects the two findings of lower inflation rates for consumption with high patenting activity, and higher patenting intensity of consumption baskets for older households. We show that inflation especially in the last years has been systematically higher for older households. This means that the baby boomer cohort experiences comparatively lower overall inflation, particularly in the latter half of the sample period, when innovation activity increasingly shifts toward goods disproportionately consumed by this cohort.

We examine how cohort-specific inflation rates accumulate over time. For each year, we compute the inflation rate faced by a given birth cohort by weighting price changes across consumption items by that cohort's expenditure basket. Applying this procedure to the first half of our sample (1984–1998), we find that the cumulative price index by 1998 does not indicate a systematic age component. This pattern is consistent with the relatively flat age gradient in patenting intensity during this period.

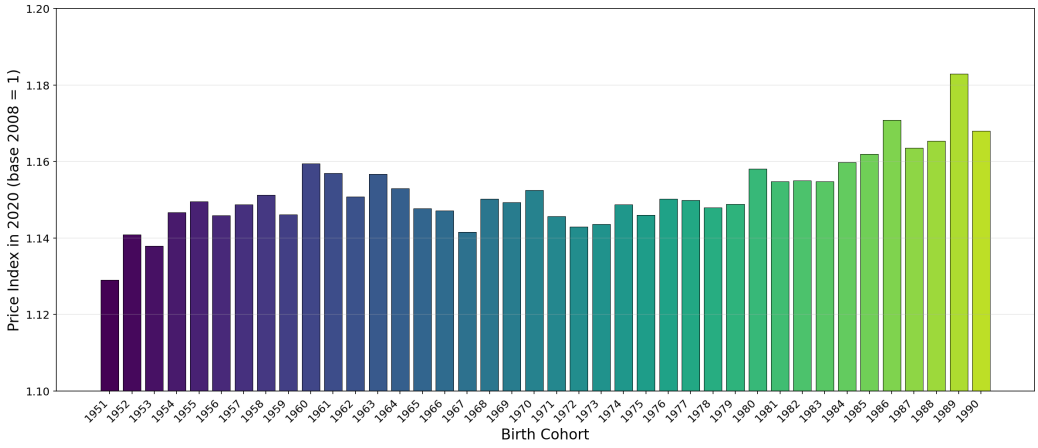
In contrast, in the second half of the sample, inflation outcomes diverge across cohorts. Younger cohorts experience systematically higher cumulative inflation, while the price index for the older baby boomer cohort evolves more favorably. Importantly, this widening inflation differential coincides with a shift in innovation activity toward goods more intensively consumed by older cohorts, suggesting a link between the direction of technological change and cohort-specific inflation dynamics.

Figure 8: Cohort specific price index

(a) 1984-1998



(b) 2008-2020



Note: Panel (a) shows the cohort specific consumption price index in 1998, based to 1 in 1984. Panel (b) shows the index for 2008-2020.

5 Mechanism

This section develops the economic mechanism that rationalizes our empirical findings. We describe the key forces through which demographic change endogenously redirects innovation and generates intergenerational consumption inequality, and we discuss the resulting inefficiency of the market allocation of R&D.

5.1 Setup

Demand. The economy is populated by overlapping generations of households that differ in age a and income y_a. Households allocate expenditure across C consumption sectors. We allow preferences to be non-homothetic, so that expenditure shares depend on both age and income:

$$\omega_{a,c} = \omega(a, y_a, c). \tag{5}$$

Non-homotheticity is important for two reasons. First, even holding age fixed, richer households spend relatively more on certain goods (financial services, travel, healthcare beyond basic needs) and less on others (food, tobacco). Second, and crucially for our mechanism, it implies that the life-cycle accumulation of wealth alters the *composition* of demand, not just its level. As households age and grow wealthier, their baskets shift toward income-elastic categories. With homothetic preferences, the income channel in $y_{a,t}$ would scale demand uniformly across sectors and have no effect on the direction of innovation. Non-homotheticity ensures that demographic aging and wealth accumulation jointly reshape the sectoral composition of demand.

In what follows, we treat the expenditure shares $\omega_{a,c}$ as a reduced-form object that captures both taste and income effects, and we calibrate them directly from the age-specific expenditure profiles documented in Section 4. These profiles are treated as fixed over time: we do not allow preferences to shift, so all time variation in the direction of innovation is driven by changes in the demographic and income structure of the population.

Two features of the empirical expenditure profiles are central to the mechanism. First, the age profile of expenditure shares is highly nonlinear. Between ages 25 and 45, expenditure shares vary little across age groups: a 30-year-old and a 45-year-old consume broadly similar baskets. The sharp divergence occurs after age 50, when healthcare spending rises steeply, education expenditure collapses, and housing costs decline with mortgage payoff. Second, the categories that are most tilted toward older households (healthcare, financial services, utilities) are also among the most patent-intensive, while the categories most tilted toward younger households (education, housing) have low

patent intensity.

Supply and innovation. Each consumption sector c produces a differentiated good at quality level $Q_{c,t}$. Quality improves through innovation. When a firm successfully innovates in sector c, quality increases by a factor $\lambda > 1$: the innovator produces at quality $\lambda Q_{c,t}$ while competitors remain at $Q_{c,t}$. This gives the innovator a temporary cost advantage that translates into profits. Under standard Bertrand competition on the quality ladder, the successful innovator captures a flow profit proportional to the market size of sector c:

$$\pi_{c,t} = \mu \cdot M_{c,t}, \tag{6}$$

where μ is a markup that depends on the innovation step size λ, and $M_{c,t}$ is the effective market size of sector c:

$$M_{c,t} = \sum_{a} N_{a,t} \cdot \omega_{a,c} \cdot y_{a,t}. \tag{7}$$

Here $N_{a,t}$ is the population of age group a at time t and $y_{a,t}$ is their total spending. Market size thus combines headcount, spending power, and the expenditure share allocated to sector c.

Innovation in sector c requires R&D investment. We assume that the probability of a successful innovation is concave in the amount of R&D effort allocated to the sector, $R_{c,t}$:

$$\iota_{c,t} = \eta \cdot R_{c,t}^{\alpha}, \quad \alpha \in (0,1), \tag{8}$$

where η is a productivity parameter and $\alpha < 1$ captures diminishing returns, reflecting congestion or duplication in research. Diminishing returns ensure an interior solution: in equilibrium, all sectors receive some innovation, but sectors with larger markets receive more.

Equilibrium allocation of R&D. With free entry into R&D, researchers allocate effort across sectors until the expected return is equalized. The expected return to a unit of R&D in sector c is the probability of success times the flow profit:

$$\frac{\partial \iota_{c,t}}{\partial R_{c,t}} \cdot \pi_{c,t} = w_t^R \quad \forall\, c, \tag{9}$$

where w_t^R is the cost of a unit of R&D effort, equalized across sectors in equilibrium. Substituting (6) and (8):

$$\alpha\eta R_{c,t}^{\alpha-1} \cdot \mu \cdot M_{c,t} = w_t^R \quad \Rightarrow \quad R_{c,t} = \left(\frac{\alpha\eta\mu \cdot M_{c,t}}{w_t^R} \right)^{\frac{1}{1-\alpha}}. \tag{10}$$

This delivers the central prediction: R&D effort in sector c is an increasing and convex function of market size $M_{c,t}$. Sectors with larger markets attract disproportionately more innovation. Total R&D is pinned down by the aggregate resource constraint, so an increase in $M_{c,t}$ for old-consumer sectors comes partly at the expense of innovation in young-consumer sectors. This reallocation, not just a level effect, is what generates the distributional consequences.

Innovation and prices. Accumulated innovation lowers quality-adjusted prices. As quality improves along the ladder, the price per unit of quality in sector c evolves as:

$$p_{c,t} = \frac{\bar{p}_c}{Q_{c,t}}, \tag{11}$$

where $\bar{p}_c$ is a cost parameter. The growth rate of quality depends on the innovation rate: $\hat{Q}_{c,t} = \iota_{c,t} \cdot \log \lambda$. Sectors receiving more R&D have faster quality growth and thus faster price decline. This is consistent with the impulse response evidence in Section 4.1, which shows that breakthrough innovation triggers follow-on patenting and a delayed but persistent decline in category-level inflation.

5.2 Why the age gradient of innovation steepens over time

A central fact documented in Section 4 is that the age gradient of patent intensity is flat in 1984–1999 but steep in 2000–2019. This time variation arises endogenously from the interaction of a fixed demographic transition with the nonlinear life-cycle structure of consumption, without requiring any change in preferences.

The patent intensity of age group a in year t is:

$$PI_{a,t} = \sum_c \omega_{a,c} \cdot P_{c,t}, \tag{12}$$

where $P_{c,t}$ is the equilibrium number of patents in sector c, determined by the R&D allocation in (10). The age gradient of patent intensity changes over time only because the supply of patents $P_{c,t}$ shifts across sectors. With fixed $\omega_{a,c}$, the entire time variation runs through the supply side.

The mechanism operates through the position of the large cohort on the nonlinear expenditure schedule. The baby boomers are always the largest cohort, but the demand signal they send to innovators depends on where they sit on the $\omega_{a,c}$ schedule. When the boomers are in their thirties (circa 1990), their expenditure shares differ little from those of neighboring age groups: the ω schedule is relatively flat in this range, so a large cohort at age 30 pulls innovation in roughly the same directions as a large cohort at age 40 would. The product $N_{boomer} \times \omega_{30,c}$ does not disproportionately favor any particular sector. When the same cohort reaches its sixties (circa 2020), however, its expenditure shares become highly distinctive: sharply tilted toward healthcare, financial services, and home maintenance, and away from education and housing. Now the product $N_{boomer} \times \omega_{60,c}$ generates a strong pull toward old-consumer sectors. The nonlinearity of $\omega_{a,c}$ in age is the key: the mapping from cohort size to market size signals is age-dependent, and it only becomes powerful once the large cohort reaches the part of the life cycle where consumption composition diverges most from that of other age groups.

Non-homotheticity amplifies the timing through the income channel. Life-cycle wealth accumulation implies that $y_{a,t}$ peaks around ages 50 to 60. With homothetic preferences, higher income would raise $M_{c,t}$ proportionally across all sectors consumed by the old, without changing the relative pull across sectors. With non-homothetic preferences, however, the wealthier boomers at age 60 spend disproportionately more on income-elastic old-consumer goods (healthcare, financial services) relative to necessities. This tilts $M_{c,t}$ further toward precisely the sectors where the age effect of $\omega_{a,c}$ is already strongest, creating a multiplicative interaction between the demographic and income channels.

A further amplification channel operates through consumer inertia. As documented by Bornstein (2025), older households exhibit greater brand loyalty and lower switching rates. As the boomers age into old-consumer sectors, the average consumer in those sectors becomes stickier. This raises the private appropriability of incremental innovation by incumbents: the effective markup μ in equation (6) is higher in sectors with inertial consumers, pulling additional R&D toward these categories beyond what pure market size would dictate. Importantly, this inertia channel specifically favors incremental quality improvements over radical product innovation, which is consistent with our finding (Figure A.8) that the age gradient of patent intensity is visible for patents held by publicly traded firms but disappears for breakthrough patents.

5.3 From directed innovation to intergenerational inequality

The redirection of innovation toward old-consumer sectors generates intergenerational consumption inequality through the quality-price channel described above.

Using (11), the growth rate of real consumption for age group a is:

$$g^U_{a,t} \approx \sum_c \omega_{a,c} \cdot \hat{Q}_{c,t} = \sum_c \omega_{a,c} \cdot \iota_{c,t} \cdot \log \lambda. \tag{13}$$

When innovation concentrates in old-consumer categories, $\iota_{c,t}$ is systematically larger for goods with high $\omega_{old,c}$, so $g^U_{old} > g^U_{young}$. This differential requires no change in nominal incomes, wages, or labor market structure. It operates entirely through the product market: even if all households earn the same nominal income growth, those whose baskets are more exposed to innovation enjoy faster growth in real consumption.

This differential accumulates over time. The cohort-specific price indices in Figure 8 confirm the pattern: the cumulative inflation differential between the youngest and oldest cohorts widens markedly in the second half of the sample, reaching several percentage points by 2020.

Importantly, the general equilibrium structure of the model implies that the reallocation of R&D toward old-consumer sectors is not merely a level effect. Because total R&D resources are scarce, equation (10) implies that an increase in $M_{c,t}$ for old-consumer sectors raises $R_{c,t}$ in those sectors but, through the equilibrium research wage w^R_t, crowds out R&D in young-consumer sectors. Innovation is thus a relative-sum game across sectors: the young do not merely fail to benefit from the demographic shift; they are actively made worse off, as the sectors they rely on, education, housing, childcare, receive less innovation than they would in the absence of the boomer demand pull.

5.4 Inefficiency of market-directed innovation

A natural question is whether the observed redirection of innovation is socially optimal. If the population ages, a benevolent planner might also choose to direct more R&D toward goods consumed by the elderly. The interesting question is whether the market *over-directs* innovation relative to the planner, and if so, through which channels.

Consider a utilitarian planner who maximizes the discounted sum of welfare across all age groups. The planner's first-order condition for allocating R&D across sectors

equates the *social* marginal return to R&D:

$$\frac{\partial \iota_{c,t}}{\partial R_{c,t}} \cdot \underbrace{\sum_a N_{a,t} \cdot u'(C_{a,t}) \cdot \omega_{a,c} \cdot C_{a,t} \cdot \left(1 + \sum_{s=1}^{T_a} \beta^s\right)}_{\text{Social value of innovation in sector } c} = \nu_t \quad \forall\, c, \tag{14}$$

where ν_t is the shadow value of R&D resources, $u'(C_{a,t})$ is the marginal utility of consumption of age group a, T_a is the remaining lifetime of age group a, and β is the discount factor. Compare this with the market first-order condition (9), which equates the *private* marginal return:

$$\frac{\partial \iota_{c,t}}{\partial R_{c,t}} \cdot \mu \cdot \sum_a N_{a,t} \cdot \omega_{a,c} \cdot C_{a,t} = w_t^R \quad \forall\, c. \tag{15}$$

The comparison reveals three distinct wedges between the market and planner allocations, all biasing the market toward old-consumer sectors.

The marginal utility wedge. The market weights demand by expenditure in dollars. The planner weights it by marginal utility $u'(C_{a,t})$, which is declining in consumption and therefore in wealth. Since older households are wealthier, $u'(C_{old,t}) < u'(C_{young,t})$: the planner places relatively more value on quality improvements in sectors consumed by the young. As wealth inequality between age groups has widened over the sample period, this wedge has grown.

The horizon wedge. The planner values a quality improvement over the consumer's remaining lifetime T_a, captured by the term $(1 + \sum_{s=1}^{T_a} \beta^s)$ in (14). A quality improvement in education, consumed primarily by 25-year-olds, generates benefits for decades. The same improvement in a healthcare product consumed by 65-year-olds benefits consumers over a shorter horizon. The market, in contrast, captures only contemporaneous profits and does not internalize the longer stream of future welfare gains from innovation in young-consumer sectors.

The appropriability wedge. The markup μ in the market first-order condition is not uniform across sectors. In sectors where older, more inertial consumers dominate, incumbents face less competitive pressure and can appropriate a larger share of the surplus from innovation. This effectively raises the private return to R&D in old-consumer sectors relative to the social return. The inertia documented by Bornstein (2025) acts as a friction that attracts R&D to sectors where firms can extract rents from

captive demand, not where the social return to innovation is highest.

The combination of these three wedges implies that the market equilibrium features systematically more innovation in old-consumer goods and less innovation in young-consumer goods than is socially optimal. The welfare cost of this misallocation is an intergenerational transfer, mediated by the innovation system, from the young to the old.

5.5 Quantitative implications and calibration targets

Our empirical results provide a rich set of moments to discipline and evaluate this mechanism quantitatively. The life-cycle expenditure profiles (Figure A.4) pin down the demand-side parameters $\omega_{a,c}$. The income elasticities embedded in these profiles, in particular the rising share of healthcare and financial services with age and income, discipline the degree of non-homotheticity. The cohort size regression (Table 2) identifies the elasticity of R&D to market size, which corresponds to $1/(1-\alpha)$ in equation (10). The innovation-price impulse responses (Figure 3) identify the speed and magnitude with which innovation translates into price reductions, pinning down the innovation step size λ and the diffusion lag. The demographic transition itself, the time path of $N_{a,t}$, is observed directly.

Given these inputs, the model generates predictions for two objects that can serve as out-of-sample validation targets. First, the age gradient of patent intensity and its evolution across subperiods (Figure 6): the model predicts a flat gradient when the boomers are young and a steep gradient when they are old. Second, the cohort-specific price indices (Figure 8): the model predicts no differential in the early sample and a widening gap in the later sample.

A natural decomposition exercise emerges. One can feed the model with the observed demographic transition alone (holding income profiles at their 1984 values) and compare the predicted age gradient and inflation differentials with those obtained when allowing income and wealth to evolve as observed. The difference measures the amplification from non-homothetic demand. Similarly, one can shut down the inertia channel by setting the markup μ to be uniform across sectors and quantify how much of the excess innovation in old-consumer sectors is attributable to appropriability rather than market size. Even if the directed innovation channel accounts for only a fraction of the total divergence in cohort-specific inflation, the finding that demographics alone can endogenously generate meaningful intergenerational consumption inequality, through a mechanism that no existing policy targets, constitutes a novel contribution.

6 Conclusion

This paper constructs a new dataset linking nearly six million U.S. utility patents to detailed consumption categories over a fifty-year period. Three main findings emerge. First, the direction of innovation has shifted systematically toward goods consumed by older households, driven by the aging of the baby boomer cohort through the nonlinear life-cycle structure of consumption. Second, breakthrough innovation generates persistent declines in category-level inflation, reaching 0.8 percentage points below baseline after nine years. Third, because innovation increasingly targets old-consumer sectors, cohort-specific inflation paths have diverged: younger cohorts face higher cumulative price growth, while the baby boomer cohort benefits from more favorable price dynamics. This divergence is absent in the first half of our sample, when the age gradient of innovation was still flat.

We rationalize these findings through a directed innovation mechanism in which the sole exogenous driver is the demographic transition. As a large cohort ages into the part of the life cycle where its consumption becomes most distinctive, the effective market size of old-consumer sectors grows, pulling R&D in its wake. The resulting quality improvements and cost reductions benefit the large older cohort disproportionately, generating intergenerational consumption inequality endogenously. We argue that the market over-directs innovation toward older consumers relative to the social optimum, because R&D responds to market size in dollars rather than to marginal utility, because it ignores the longer remaining lifetime of young consumers, and because consumer inertia among the elderly raises private appropriability without corresponding social gains.

Our findings have implications for both measurement and policy. On measurement, standard aggregate price indices mask economically meaningful variation in the inflation rates experienced by different cohorts, and this variation is partly endogenous to the innovation system. On policy, the results suggest that the welfare costs of demographic aging extend beyond the fiscal pressures on pensions and healthcare that dominate the public debate. By redirecting the innovation effort of the economy, aging also raises the relative cost of goods essential to the young, including education, housing, and childcare, through a channel that no existing policy instrument directly addresses. Understanding the quantitative magnitude of this channel, and designing policies that could correct the resulting misallocation of R&D, is an important direction for future work.

References

Acemoglu, D. (1998). Why do new technologies complement skills? Directed technical change and wage inequality. *Quarterly Journal of Economics*, 113(4):1055–1089.

Acemoglu, D. (2002a). Directed technical change. *Review of Economic Studies*, 69(4):781–809.

Acemoglu, D. (2002b). Technical change, inequality, and the labor market. *Journal of Economic Literature*, 40(1):7–72.

Acemoglu, D., Akcigit, U., and Celik, M. A. (2022). Radical and incremental innovation: The roles of firms, managers, and innovators. *American Economic Journal: Macroeconomics*, 14(3):199–249.

Acemoglu, D. and Linn, J. (2004). Market size in innovation: Theory and evidence from the pharmaceutical industry. *Quarterly Journal of Economics*, 119(3):1049–1090.

Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.

Acemoglu, D. and Restrepo, P. (2022). Demographics and automation. *Review of Economic Studies*, 89(1):1–44.

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., and Hémous, D. (2019). Innovation and top income inequality. *Review of Economic Studies*, 86(1):1–45.

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2023). A theory of falling growth and rising rents. *Review of Economic Studies*, 90(6):2617–2652.

Aghion, P., Bergeaud, A., Lequien, M., and Melitz, M. J. (2024). The heterogeneous impact of market size on innovation: Evidence from french firm-level exports. *Review of Economics and Statistics*, 106(3):608–626.

Aguiar, M. and Hurst, E. (2013). Deconstructing life cycle expenditure. *Journal of Political Economy*, 121(3):437–492.

Akcigit, U. and Ates, S. T. (2021). Ten facts about the business dynamism of the U.S. economy. *American Economic Journal: Macroeconomics*, 13(1):257–298.

Aksoy, Y., Basso, H. S., Smith, R. P., and Grasl, T. (2019). Demographic structure and macroeconomic trends. *American Economic Journal: Macroeconomics*, 11(1):193–222.

Ameriks, J. and Zeldes, S. P. (2002). How do household portfolio shares vary with age? TIAA-CREF Workin Paper 112(2).

Argente, D., Baslandze, S., Hanley, D., and Moreira, S. (2020). Patents to products: Product innovation and firm dynamics. Working Paper.

Arvai, K. and Mann, K. (2026). Consumption inequality in the digital age. unpublished manuscript.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics*, 135(2):645–709.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333.

Bekamiri, H., Hain, D. S., and Jurowetzki, R. (2021). Patentsberta: A deep nlp based hybrid model for patent distance and classification using augmented sbert. *arXiv preprint arXiv:2103.11933*.

Bergeaud, A., Potiron, Y., and Raimbault, J. (2017). Classifying patents based on their semantic content. *PLoS ONE*, 12(4):e0176310.

Bloom, N., Hassan, T. A., Kalyani, A., Lerner, J., and Tahoun, A. (2021). The diffusion of disruptive technologies. Working Paper 28999, NBER.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144.

Bornstein, G. (2025). Entry and profits in an aging economy: The role of consumer inertia. Working Paper 33820, NBER.

Clemens, J. (2013). Medicare and the rise of american medical patenting: The economics of user-driven innovation. Working Paper, University of California San Diego.

Deaton, A. S. and Paxson, C. (1994). Saving, growth, and aging in Taiwan. In Wise, D., editor, *Studies in the Economics of Aging*, pages 331–362. University of Chicago Press.

Dyevre, A. and Seager, O. (2023). 70 years of patents matched to compustat firms: Methodology and insights about firm heterogeneity. unpublished working paper.

Fagereng, A., Gottlieb, C., and Guiso, L. (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance*, 72(2):705–750.

Fernández-Villaverde, J. and Krueger, D. (2007). Consumption over the life cycle: Facts from consumer expenditure survey data. *Review of Economics and Statistics*, 89(3):552–565.

Finkelstein, A. (2004). Static and dynamic effects of health policy: Evidence from the vaccine industry. *Quarterly Journal of Economics*, 119(2):527–564.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1):47–89.

Hobijn, B. and Lagakos, D. (2009). Inflation inequality in the United States. *Review of Income and Wealth*, 55(3):540–567.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the US retail sector. *Quarterly Journal of Economics*, 134(2):715–783.

Jaravel, X. and Lashkari, D. (2024). Measuring growth in consumer welfare with income-dependent preferences: Nonparametric methods and estimates for the United States. *Quarterly Journal of Economics*, 139(1):477–532.

Jones, B. F. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Review of Economic Studies*, 76(1):283–317.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182.

Kaplan, G. and Schulhofer-Wohl, S. (2017). Microeconomic heterogeneity and macroeconomic shocks. *Journal of Economic Perspectives*, 31(2):167–192.

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2021). Measuring technological innovation over the long run. *American Economic Review: Insights*, 3(3):303–320.

Kogan, L., Papanikolaou, D., Schmidt, L. D., and Seegmiller, B. (2023). Technology and labor displacement: Evidence from linking patents with worker-level data. Working Paper w31846, National Bureau of Economic Research.

Lybbert, T. J. and Zolas, N. J. (2014). Getting patents and economic data to speak to each other: An "algorithmic links with probabilities" approach for joint analyses of patenting and economic activity. *Research Policy*, 43(3):530–542.

Maestas, N., Mullen, K. J., and Powell, D. (2023). The effect of population aging on economic growth, the labor force, and productivity. *American Economic Journal: Macroeconomics*, 15(2):306–332.

Mann, K. and Püttmann, L. (2023). Benign effects of automation: New evidence from patent texts. *Review of Economics and Statistics*, 105(3):580–595.

Schmookler, J. (1966). *Invention and Economic Growth*. Harvard University Press, Cambridge, MA.

Appendix

A Additional Figures and Tables

Figure A.1: Total vs. matched patents over time

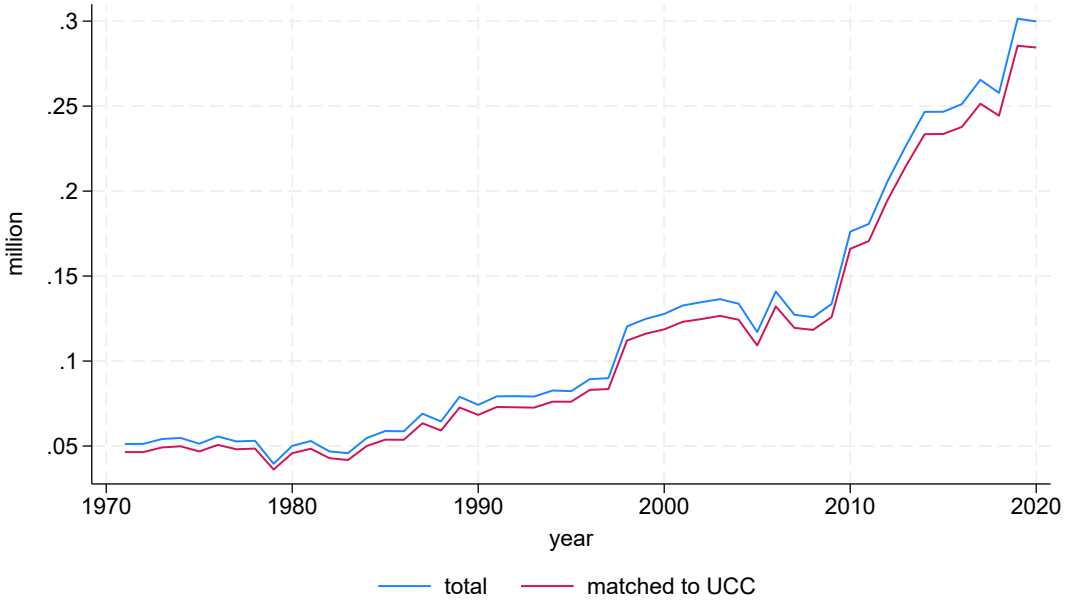


Figure A.2: Share of patents matched to at least one UCC

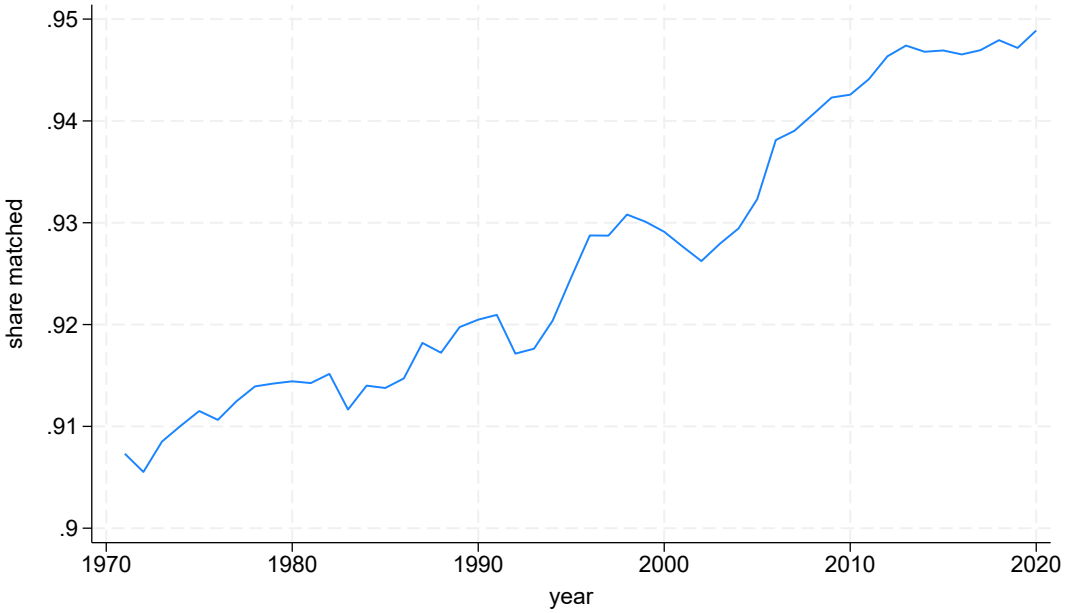


Figure A.3: Average number of matched UCCs

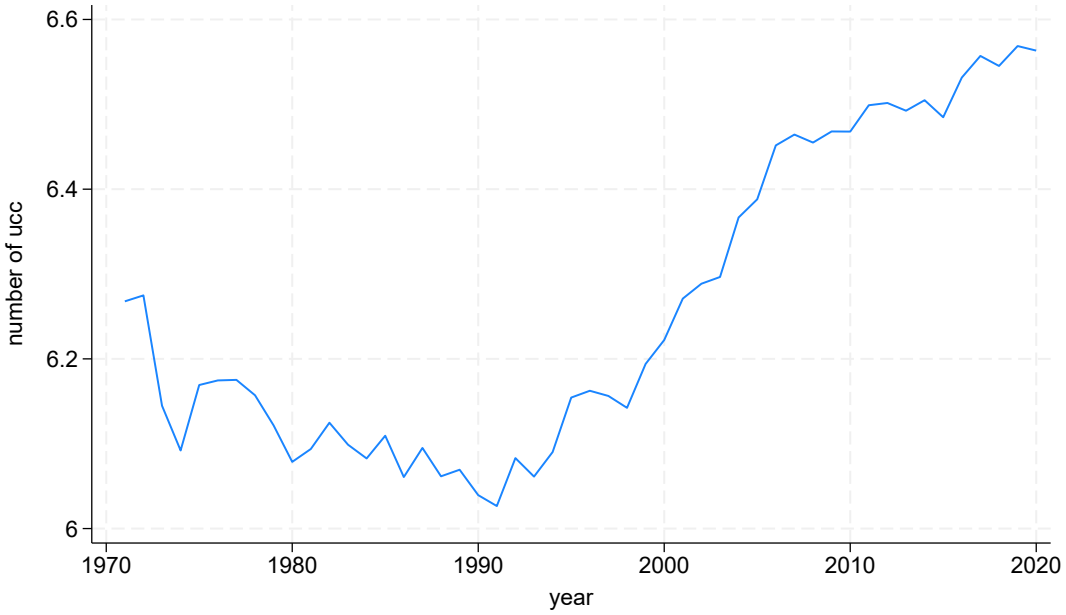


Figure A.4: Life-cycle consumption profiles

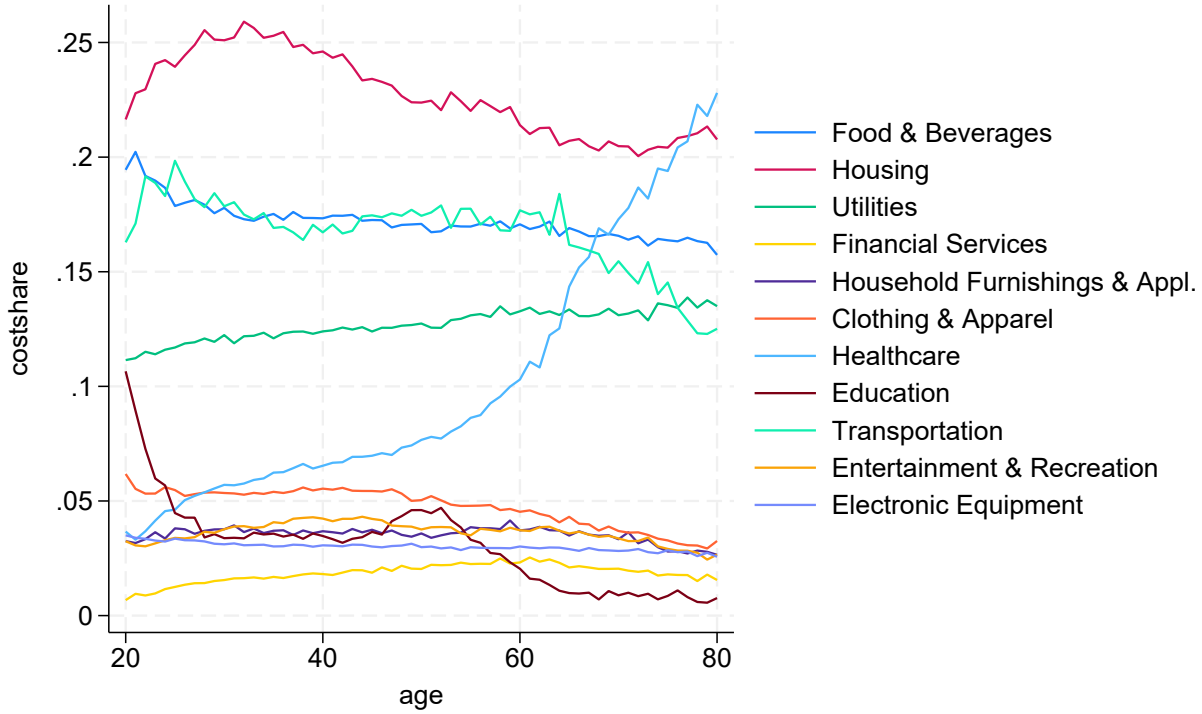


Table A.1: UCCs and supercategories

supercat	ucc_const	ucc
1	190901-200900, 790220-790430, 800700	190901-200900, 790220-790430, 800700
2	210110-220311, 230901, 790600, 800710	210110-220322, 230901-230902, 350110, 790600, 800710, 880110-880310
3	230112-230115, 230121-230150, 240111-240321, 320410-320420, 320522-320632, 320902, 330511, 340310-340420, 340911-340914, 990920	230111-230116, 230119-230152, 240111-240323, 320161-320163, 320410-320420, 320522-320633, 320902, 330511, 340310-340420, 340520-340530, 340901, 340903, 340911-340915, 690241-690242, 790690, 990920-990940
4	250111-270211, 270411-270901, 320232, 470111-470211	250111-270214, 270411-270904, 320232, 470111-470212, 690210
5	2120, 680902, 700110-710110	2120-5620, 620112, 680220, 680902, 700110-710110
6	280110-290440, 320110-320120, 320220, 320233-320345, 320901, 320903, 420110	280110-290440, 320110-320120, 320210-320231, 320233-320360, 320901, 320903, 340630, 340904, 420110-420120
7	360110-410901, 430110-440110, 440130, 440900	360110-410904, 430110-430120, 440110, 440130-440150, 440900
8	340906, 340910, 540000-580903, 640420-650310	340906, 340910, 540000-580910, 640420-650900
9	340211, 530902, 590220, 660110-670902	340210-340212, 530902, 590220-590230, 660001-670903
10	220901, 340510, 450110-460902, 470220-530901, 600110-600142, 600310, 620909	220901-220902, 340510, 450110-460902, 470220-480213, 480216-490501, 490900-530901, 600110-600142, 600310-600311, 620113-620114, 620902, 620906-620907, 620909, 620919-620922, 850300
11	310400, 320130-320150, 320904, 590310, 600210, 600410-620903, 620926-620930	310400, 320130-320150, 320904, 430130, 590110-590212, 590310-590410, 600210, 600410-620111, 620115-620420, 620903-620905, 620908, 620926-620930, 680310-680320, 680904-680905
12	630110-630210	630110-630210
13	230117, 300111-300411, 320370, 320511-320521, 440120, 440210, 690244	230117-230118, 300111-300412, 320370, 320511-320521, 340620, 340907, 440120, 440210, 690243-690245, 990900
14	270310, 310110-310350, 340908, 690111	270310-270311, 310110-310350, 340610, 340902, 340905, 340908, 480214-480215, 490502, 620912-620918, 690111-690120, 690310-690350
15	640130, 680110-680901, 690230	640130, 680110-680210, 680901, 690220-690230

Table A.2: Patents with 40 or more UCC matches

patent	title	main CPC	assignee	grant year	UCCs
9288053	Schema signing	cryptographic mechanisms for authentication between devices	Microsoft Technology Licensing LCC	2016	45
8244634	Interchange categories	banking, e.g. interest calculation or account maintenance	Redbridge USA Inc.	2012	45
8108260	System and method for dynamic categorization	electronic shopping	Etsy Inc.	2012	45
6148290	Service contract for managing service systems	buying, selling or leasing transactions	IBM Corp.	2000	43
7050990	Accelerated task execution	sound input and output	Apple Inc.	2022	42
8175938	Method and system for facilitating merchant-initiated online payments	establishing or using transaction specific rules	PayPal Inc.	2012	42
9396491	System, method, and computer-readable storage device for providing a buy option in search results when user input is classified as having a purchase intent	electronic shopping	Monticello Enterprises LLc	2016	41
8533208	System and method for topic extraction and opinion mining	document management systems	eBay Inc.	2013	41
7603367	Method and system for displaying attributes of items organized in a searchable hierarchical structure	database structures and file systems for information retrieval	Amazon Technologies	2009	41
9640065	Electronic apparatus, external apparatus, and method of controlling power supply to external apparatus	transmitting signals using a wireless electronic link	Samsung Electronics Co. Ltd	2017	40
8190486	Techniques for product selection	marketing, price estimation or determination	Engage3 LLC	2012	40
6910017	Inventory and price decision support	resources, workflows, human or project man-	Oracle International Corp.	2005	40

Figure A.5: Breakthrough-patent share in major categorizes over time

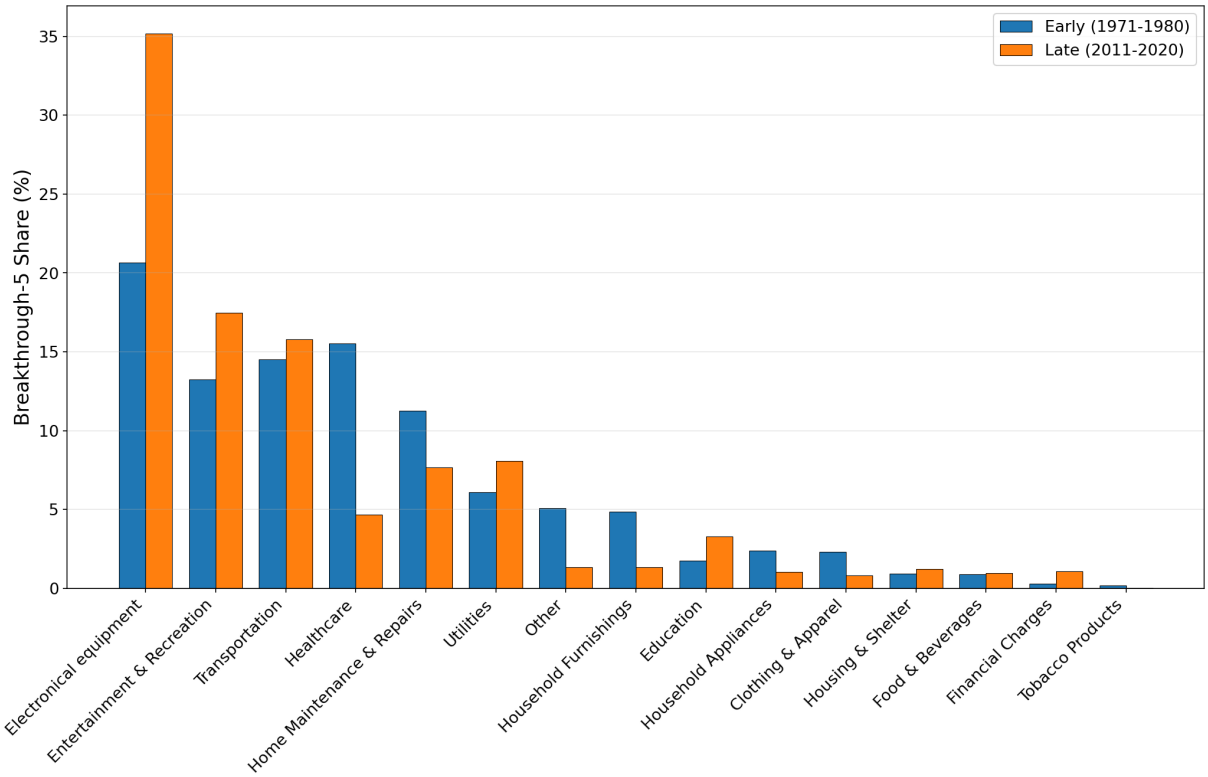
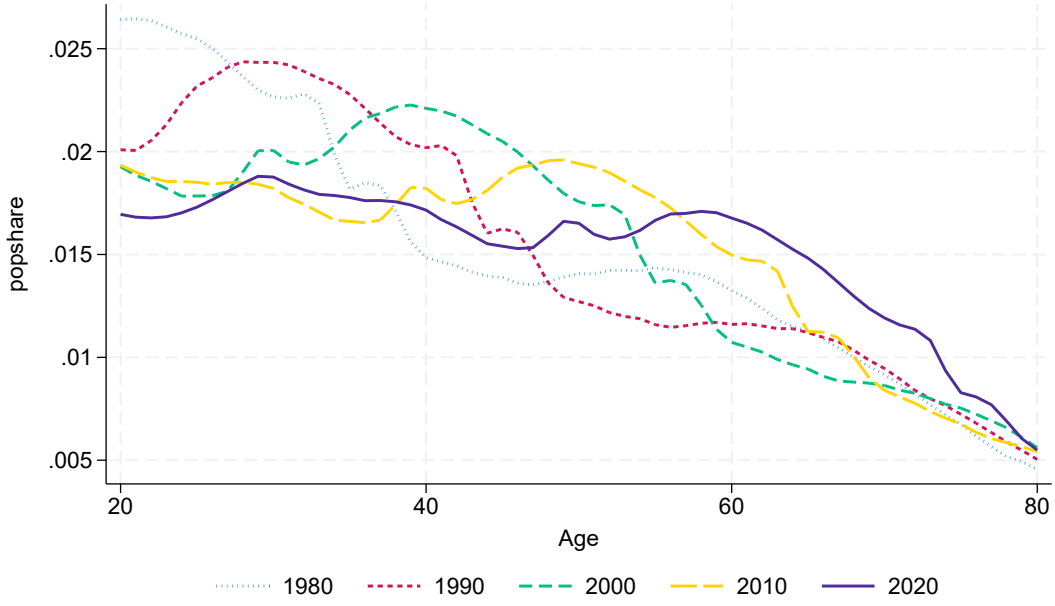
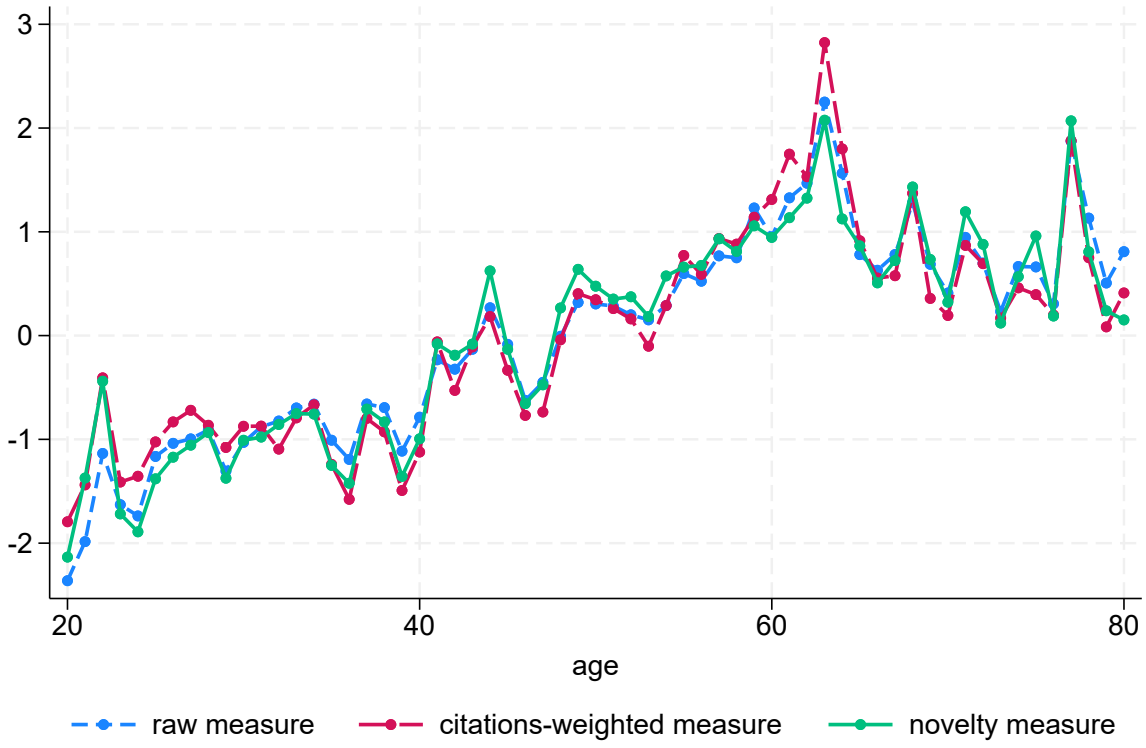


Figure A.6: Cohort sizes



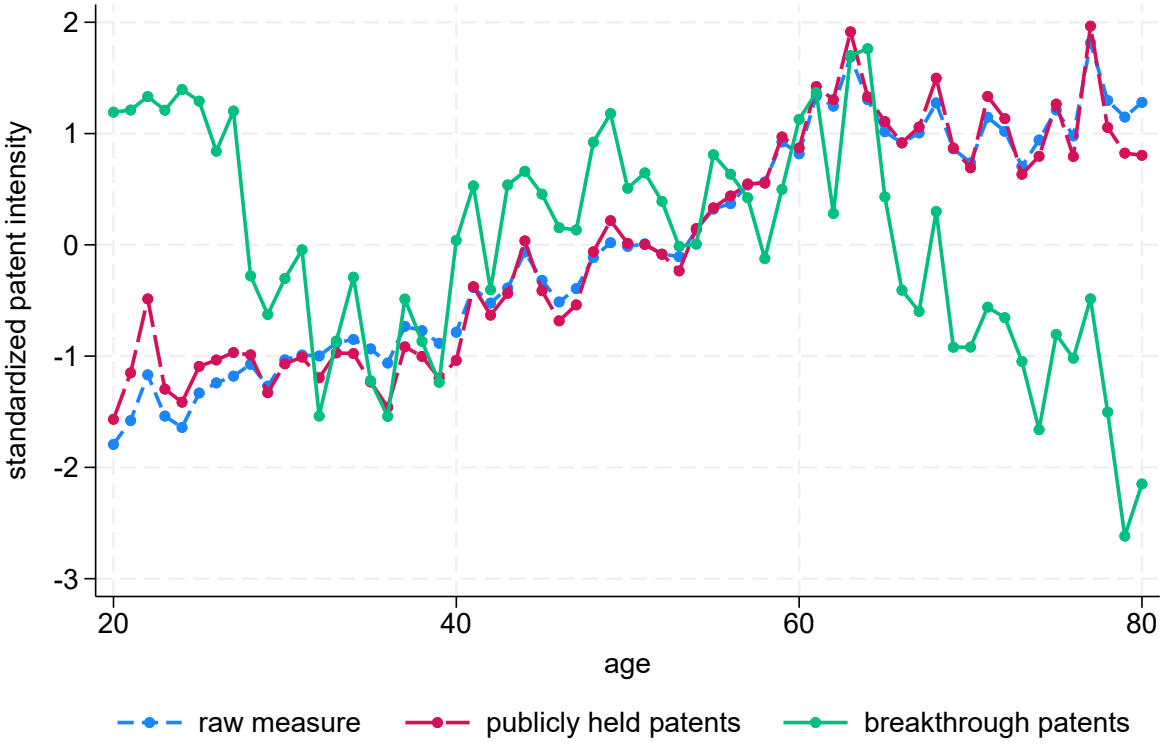
Source: United Nations World Population Prospects.

Figure A.7: Patents across age, residualized



Note: The figure shows the same data as Figure 5, but instead of showing the raw (standardized) patent intensities, it shows (standardized) residuals from a regression of patent intensity on year fixed effects, household income, education and the urban share.

Figure A.8: Patents across age, residualized



Note: The figure shows the same data as Figure 5, but comparing the raw (standardized) patent measure with one measure constructed only with publicly held patents, and one constructed only with breakthrough pattents as defined by Kelly et al. (2021).

Figure A.9: Patent intensity across cohorts - linear time trend

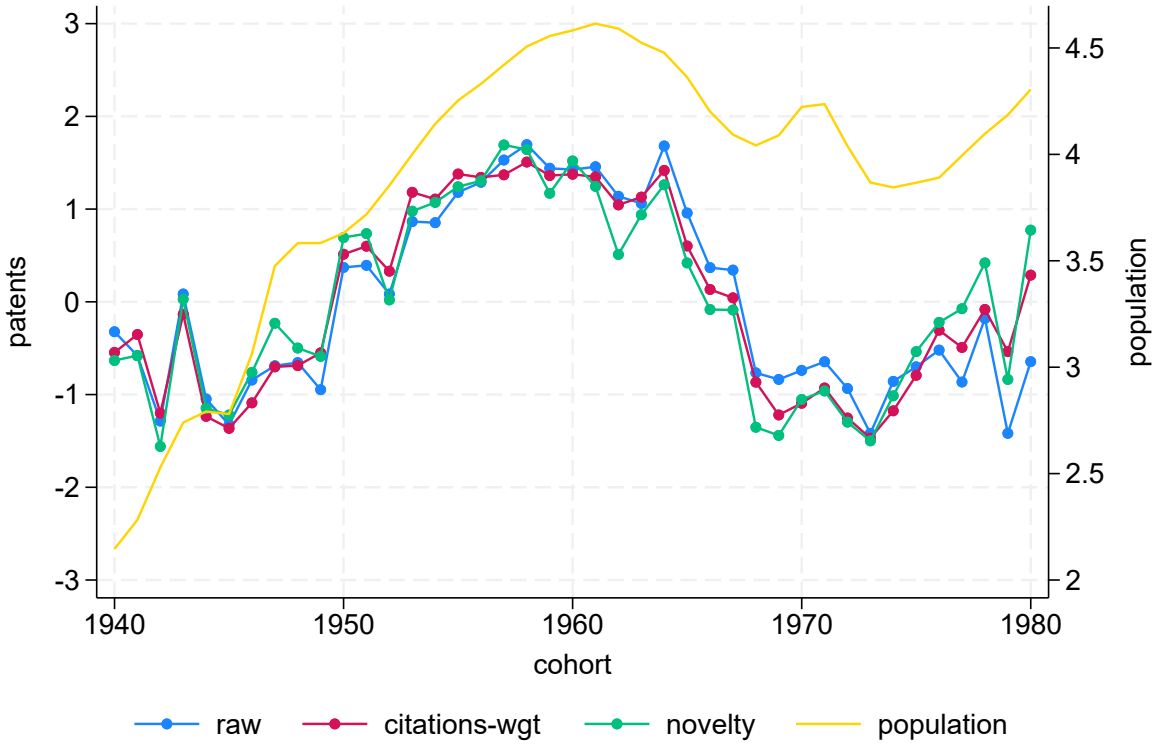
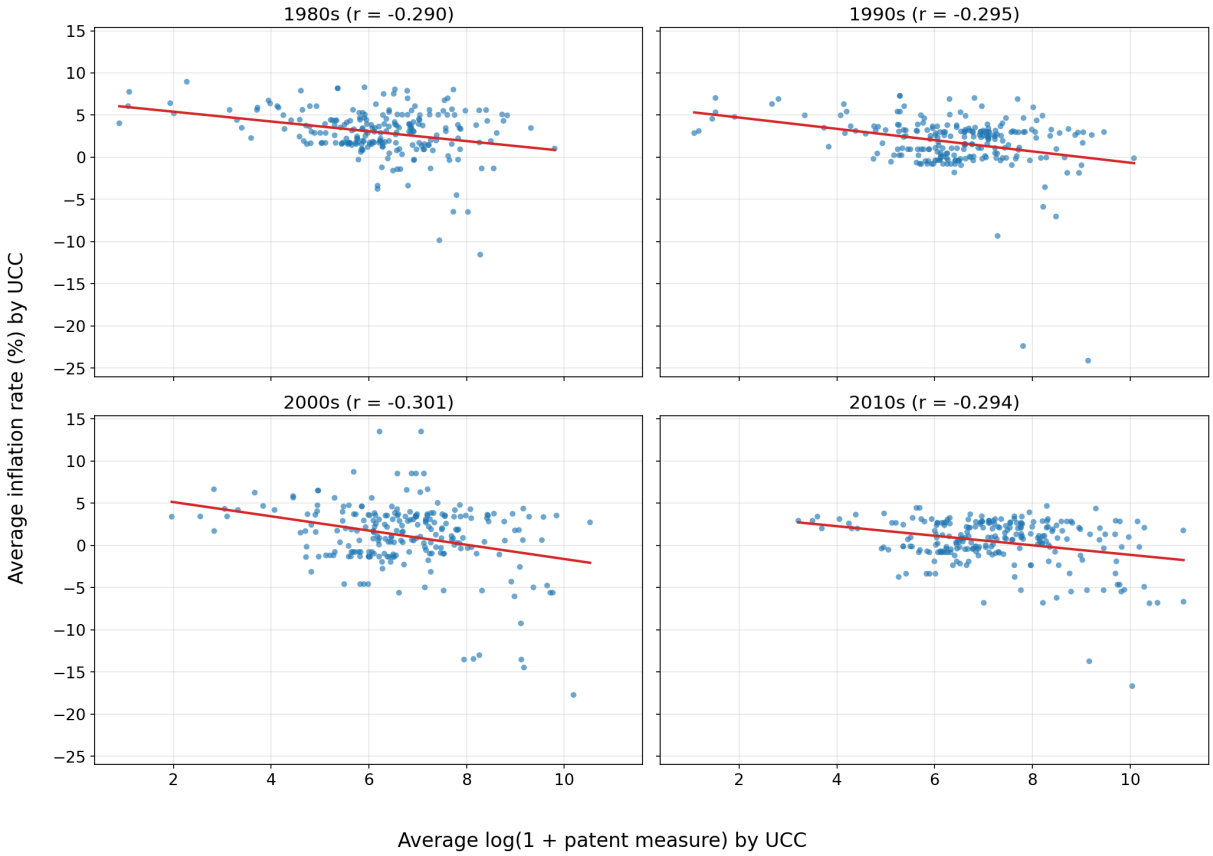


Figure A.10: Correlation average inflation and patents



Note: This panel shows the correlation between average inflation rate of a ucc consumption category and its average log patent level per decade. The red line indicates the slope of the average correlation.

B Data sources

B.1 Survey of Consumer Expenditures (CEX)

Data for 1984-2020

The Survey of Consumer Expenditures is the most detailed expenditure survey in the United States, carried out at the household level. Next to providing detailed expenditure data, it also contains a host of demographic and income information on the households. We use the interview survey data of the public-use microdata, which covers the full consumption basket from 1990 to 2023. Sample households are interviewed at a quarterly frequency for four to five consecutive months. We treat each interview round as independent observation. Households report monthly expenditures at the universal classification code (UCC) level for the three months prior to the interview month. Additionally, the data contain quarterly expenses at the level of summary product categories.

We measure income as total before-tax annual household income. Income is self-reported prior to 2004 (FINCBTAX) and imputed from 2004 onwards (FINCBTXM). We only consider households with strictly positive income. The main income component is labor income, which accounts for xxx percent of total income. Next to labor income, the income measure includes farm and non-farm business income, social security income, interest on savings accounts or bonds, income from dividends, royalties, estates and trusts and rental income. We only consider observations with positive total household income.

We calculate annual expenditures at the level of calendar years. As survey weights refer to the interview year, we need to adjust survey weights for expenditures that relate to the previous year. This concerns households interviewed in January through March. E.g. for a household interviewed in February, the reported expenditures refer to November and December of the previous year, as well as January of the interview year. In the adjustment we follow the instructions published in the Public Use Microdata Users' Documentation.

Before 1990, the CEX only contains spending on 32 summary product categories. We follow Jaravel and Lashkari (2024) to impute UCC-level expenditures. They use microdata for 1990 and assume that the expenditure shares for any given income percentile within each of the summary product categories are the same in 1984-1989 as in 1990.

We restrict the expenditure data to only include the UCCs that appear in the annual hierarchical grouping auxiliary files provided by the BLS, as these are the categories

used for the summary tables.[1] The files are available since 1997. For earlier years, we follow Jaravel and Lashkari (2024) and apply the UCC restrictions as specified in the 1997 file. We drop the UCCs referring to pension and social security contributions and cash contributions. We are only using the interview survey, whereas the summary tables published by the BLS use both the interview and the diary survey, which contains more detailed expenditure on day-to-day purchases. Therefore, the total expenditures in our dataset do not exactly match those in the summary tables published by the BLS. In order to correct for this, we calculate a scaling factor specific to each product category and income quintile, by which we adjust our expenditure data, again following the procedure in Jaravel and Lashkari (2024).

Data for 1972-1984

Prior to 1984, no detailed microdata are available. For 1972, the BLS has published summary tables for 12 different groups of before-tax income and at the level of 42 summary product categories. We match the summary product categories to the 32 categories from the more recent data and the income groups into deciles. We again rely on the microdata for 1990 and assume that the expenditure shares for any given income decile within each of the summary product categories are the same in 1972-1983 as in 1990. Finally, we linearly interpolate expenditure shares for each group and income decile between 1972 and 1984.

B.2 Product categories

We have altogether 251 product categories (UCCs) in our final dataset. This number is considerably lower than the total number of UCCs in the CEX as we undertake several steps to condense and streamline the set of UCCs:

- Where one category has been superseded by another one with the (almost) identical content, we merge these into one. E.g. 480211 and 480213 both describe vehicle parts, equipment and accessories.
- Where categories were merged over time or dismerged, we keep the more aggregate category. E.g. water sports and miscellaneous sports equipment (600900) was divided into water sports equipment (600901) and other sports equipment (600902). We label all three UCCs as 600900. Sweaters and vests (360430) and shirts

[1] Summary tables can be found on: https://www.bls.gov/cex/tables.htm

(360410) were summarized into shirts, sweaters and vests (360420). We label all three UCCs as 360420.

- We merge UCCs where there have been technical developments within the same product category, e.g. we merge rental of video cassettes, tapes, films and discs (620912) and rental of video game hardware an software (620916), (620917), (620918).
- Installations are merged to the corresponding product category, e.g. installation of televisions (690320) is merged to televisions (310140).

We need to include both a mapping from current to constant uccs and a table that lists the constant uccs and the supercategory

Table B.1: UCCs and supercategories

supercat	ucc_const	ucc
1	190901-200900, 790220-790430, 800700	190901-200900, 790220-790430, 800700
2	210110-220311, 230901, 790600, 800710	210110-220322, 230901-230902, 350110, 790600, 800710, 880110-880310
3	230112-230115, 230121-230150, 240111-240321, 320410-320420, 320522-320632, 320902, 330511, 340310-340420, 340911-340914, 990920	230111-230116, 230119-230152, 240111-240323, 320161-320163, 320410-320420, 320522-320633, 320902, 330511, 340310-340420, 340520-340530, 340901, 340903, 340911-340915, 690241-690242, 790690, 990920-990940
4	250111-270211, 270411-270901, 320232, 470111-470211	250111-270214, 270411-270904, 320232, 470111-470212, 690210
5	2120, 680902, 700110-710110	2120-5620, 620112, 680220, 680902, 700110-710110
6	280110-290440, 320110-320120, 320220, 320233-320345, 320901, 320903, 420110	280110-290440, 320110-320120, 320210-320231, 320233-320360, 320901, 320903, 340630, 340904, 420110-420120
7	360110-410901, 430110-440110, 440130, 440900	360110-410904, 430110-430120, 440110, 440130-440150, 440900
8	340906, 340910, 540000-580903, 640420-650310	340906, 340910, 540000-580910, 640420-650900
9	340211, 530902, 590220, 660110-670902	340210-340212, 530902, 590220-590230, 660001-670903
10	220901, 340510, 450110-460902, 470220-530901, 600110-600142, 600310, 620909	220901-220902, 340510, 450110-460902, 470220-480213, 480216-490501, 490900-530901, 600110-600142, 600310-600311, 620113-620114, 620902, 620906-620907, 620909, 620919-620922, 850300
11	310400, 320130-320150, 320904, 590310, 600210, 600410-620903, 620926-620930	310400, 320130-320150, 320904, 430130, 590110-590212, 590310-590410, 600210, 600410-620111, 620115-620420, 620903-620905, 620908, 620926-620930, 680310-680320, 680904-680905
12	630110-630210	630110-630210
13	230117, 300111-300411, 320370, 320511-320521, 440120, 440210, 690244	230117-230118, 300111-300412, 320370, 320511-320521, 340620, 340907, 440120, 440210, 690243-690245, 990900
14	270310, 310110-310350, 340908, 690111	270310-270311, 310110-310350, 340610, 340902, 340905, 340908, 480214-480215, 490502, 620912-620918, 690111-690120, 690310-690350
15	640130, 680110-680901, 690230	640130, 680110-680210, 680901, 690220-690230

B.3 Price data

We get data on the prices of consumption goods from NIPA. We use Tables 2.4.3U , 2.4.4U and 2.4.5U. The product codes used by NIPA are matched to UCCs using mappings provided by the BLS. Where several NIPA prices are matched to one UCC, we create Tornqvist price indices at the UCC level. In the same way, we create prices at the level of supercategories.

C Code template

```
project_root/
  data/
    raw/
    intermediate/
    final/
  models/
    bert_model/
  notebooks/
    01_explore_data.ipynb
    02_quality_check.ipynb
  code/
    A0_Master
    A0prep_a_ucc_desc
    A0prep_b_localAI_selection
    A0prep_c_FinalAI_selection
    A1_EvaluateAI
    A2_merge_ucc_year_patents
    A3_merge_ucc_year_patents_CEX
    A4_analysis
  output/
    results/
    figures/
  requirements.txt
```

more details

- A0_Master. sets all the paths and runs all the code
- A0prep_a_ucc_desc
  - Inputs: all used uccs
  - Use chat gpt to create a large lists of decriptions
  - Output: ucc_descriptions
- A0prep_b_localAI_selection
  - Inputs: ucc, ucc-descriptions,patent-id, patent-abstract
  - Runs the local AI and narrows down to possibly 50 ucc categories
  - Output: A0_a_localAI_selection. A dataset that contains every patent AI and a list with 50 ucc categories.
- A0prep_c_FinalAI_selection
  - Inputs: A0_a_localAI_selection
  - Run the chat GPT API that asks chat gpt to categorize each of the patent ids given the descriptions and uccs.
  - Output: A0_a_finalAI_selection. A dataset that contains every patent-id and its corresponding match to the uccs.
- A1_EvaluateAI
  - Inputs: A0_a_finalAI_selection
  - An RA evaluates the proposed matches of the AI, marking wrong matches by the AI (type 1 error) and adding some missed matches (type II error).
  - output: graphs and tables that give an idea how well the AI is doing.
- A2_merge_ucc_year_patents
  - Inputs: A0_a_finalAI_selection, all patents with year and number of citations, all ucc (and their year if not consistent)
  - merge the patent-id and ucc selection with patent-id, their year of publication, their year of filing, and the number of citations.
  - sum for each year and each ucc, the number of patents a ucc has. plain numbers, citation weighted (perhaps different weighting method?)
  - Output: A2_ucc_year_patents. ucc, year, number of patents matched, citation weighted patents.

- A3_merge_ucc_year_patents_CEX
  - Inputs: A2_ucc_year_patents, CEX_data
  - merge both dataset using ucc and year as a common variable. dataset has then the form: year, household-id (or maybe already income percentile), ucc, expenditure (share) for ucc in a given year, number of patents for the ucc in a given year
  - compute then for each household (or income percentile) in every year the innovation content of the basket as an expenditure weighted average of all uccs times the number of patents.
  - Output: A3_ucc_year_patents_CEX. year inc percentile, innovation content, ucc, exp share, number of patents.
- A4_analysis
  - Input: A3_ucc_year_patents_CEX
  - Create some stylized facts, tables and more sophisticated analysis with this dataset
  - Output: graphs and tables.