The Consumer-Product Characteristic Project*

Laurence Ales **Rebecca Lessem** **Christopher I. Telmer**
Ariel Zetlin-Jones
Tepper School of Business, Carnegie Mellon University

This Version: June 27, 2026

Abstract

Economists often study consumption through coarse product categories such as durables, non-durables, and services. We develop a framework that instead represents goods and services in terms of a low-dimensional set of measurable product characteristics. The framework is a consumption analog of O*NET: it maps a high-dimensional product space into a structured characteristic space that can be used to compare products, organize expenditure categories, and study changes in demand. We implement the framework for 69 aggregate consumption categories from the Bureau of Economic Analysis, assigning each product a 23-dimensional vector of characteristics. We use our framework to study two episodes in U.S. consumption. During the COVID-19 pandemic, the traditional durables/non-durables/services taxonomy does not account well for the cross-product pattern of expenditure collapse and recovery. Characteristics that measure the degree of human interaction required for consumption provide a more informative explanation. Over the longer period from 1959 to 2023, the movement of expenditure toward services masks substantial heterogeneity within services. Services that require greater time investment in use experienced stronger expenditure growth. These applications show that product characteristics capture economically important variation in consumption behavior that is obscured by standard expenditure classifications.

JEL Codes: *D12, E21, C81*
Keywords: *Consumer product characteristics; consumption taxonomy; services demand.*

*E-mail: ales@cmu.edu (Ales), rlessem@andrew.cmu.edu (Lessem), chris.telmer@cmu.edu (Telmer), azj@andrew.cmu.edu (Zetlin-Jones). We thank V.V. Chari, Jeff Galak, Simon Mongey, Marla Ripoll, Todd Schoellman, Gianluca Violante, and seminar participants at the Federal Reserve Bank of Minneapolis, the University of Virginia, and the 2025 SED in Copenhagen for helpful suggestions and discussions. This research is supported by grants from CMU's Block Center for Technology and Society and by a grant from the Richard King Mellon Foundation. Errors and dubious conclusions are our own.

1 Introduction

A class of differentiated products is completely described by a vector of objectively measured characteristics.
Rosen (1974)

The analysis of consumer behavior typically takes one of two approaches. It either analyzes very specific goods and services in isolation (e.g., automobiles, housing, education, health care), or it analyzes broad categories such as durables, non-durables and services. In this paper we develop a middle ground between fine product-level detail and broad expenditure categories. We develop a framework in which individual goods and services are represented in terms of an underlying vector of *characteristics*. This results in a *taxonomy* for goods and services that goes far beyond the traditional taxonomy of durables, non-durables and services, while at the same time maintaining sufficient parsimony so as to be useful for the macroeconomic study of consumption behavior. We use our taxonomy in a number of empirical applications intended to ascertain its explanatory capabilities.

The central claim is that many economic questions about consumption are difficult to answer with the standard product categories because those categories do not describe what products are like or how they are consumed. A characteristic-based representation makes these dimensions explicit. It allows us to ask whether products that share particular features such as durability, human interaction, conspicuousness, or time intensity move together in response to shocks or over long horizons, even when they belong to different BEA categories.

A useful reference point for our framework is the O*NET, the database of occupational characteristics sponsored by the U.S. Department of Labor. The O*NET represents occupations through a structured set of worker and job characteristics; our framework applies the same idea to consumer goods and services. The occupation 'Emergency Medicine Physician' for example, is represented by characteristics that include 'Level of Biology Education,' aptitude for 'Judgment and Decision Making,' and a measure of the level of 'Concern for Others.' Our framework is analogous. We represent the good 'Sporting Equipment,' for example, by characteristics that include 'Brandedness,' 'Frequency of Purchase,' and 'Time Investment Required for Use.' By doing so we represent the vast number of goods and services that are available to the consumer in terms of a much smaller set of fundamental *characteristics*, in particular those that play an important role in the consumer's choice problem. Like O*NET, these characteristics should both provide a parsimonious description of what a particular good *is*, as well as how the good is similar (or dissimilar) to other goods.

To motivate our approach, consider the traditional taxonomy of durables, non-durables, and services during the COVID-19 pandemic. Figure 1 shows that this taxonomy broke down in terms of accounting for consumption behavior in a downturn (see also Cirelli and Gertler (2025)). Expenditure on durable goods, having been strongly pro-cyclical in the past, behaved in the opposite manner throughout much of 2020. Services collapsed, in spite of being relatively stable through

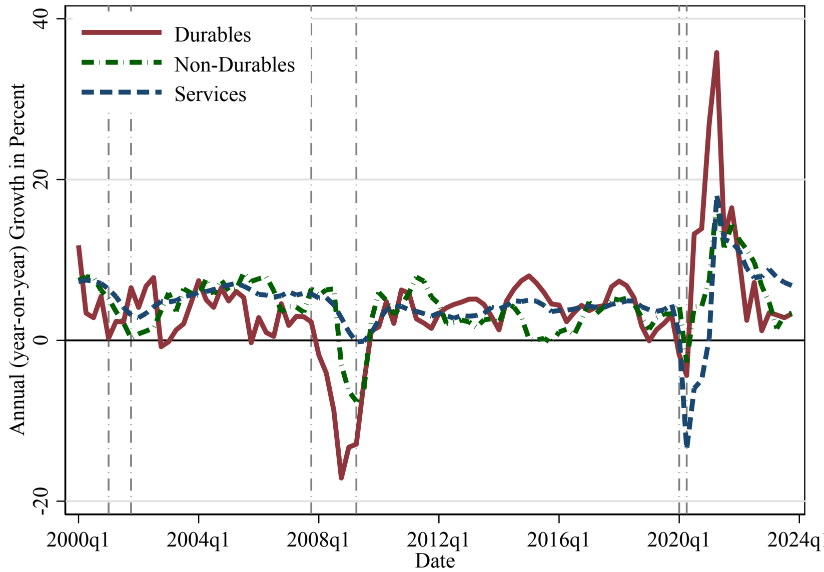


Figure 1: Growth rates of personal consumption expenditures aggregated in Durables, Non-Durables and Services, 2000–2024. Source: Bureau of Economic Analysis.

previous expansions and downturns. We also saw disparate behavior *within* the expenditure categories. Spending on durables such as automobiles fell, but it increased for other durables such as furniture. Likewise, some services behaved like services, whereas others behaved like durables. 'Communication Services,' for example, was largely unaffected, whereas 'Meals in Restaurants' and 'Accommodation' collapsed. Our objective is to develop a better taxonomy, one that can both explain Figure 1 and help us understand what might happen in the face of future macroeconomic shocks.

Figure 2 takes a longer time perspective. It shows expenditure shares and prices for several services. The left panel reports expenditure shares. We see relatively stable expenditure shares for both *Telecommunication Services* and for *Personal Care Services*, but an increase in expenditures for *Accommodation in Hotels*. Juxtaposing this against price behavior makes it obvious that not all services are alike. The large increase in the price of *Personal Care Services*, viewed alongside their stable expenditure share, suggests an elasticity of substitution of one. On the other hand *Accommodation in Hotels* saw similar price behavior but different expenditure behavior, pointing at different degrees of complementarity and substitutability. Our objective is to develop a taxonomy for goods and services that can explain, among other things, why these particular services have behaved so differently over time.

We proceed as follows. First, we develop a simple theoretical framework that blends the traditional model of consumer choice with the framework of Lancaster (1966). In our framework, consumers purchase *products* (defined to include both goods and services) subject to a budget constraint. Utility, however, is not defined over products. Rather, it is defined over a lower-dimensional set of objects that we call *consumption outcomes*. By purchasing the products 'Bicycle,'

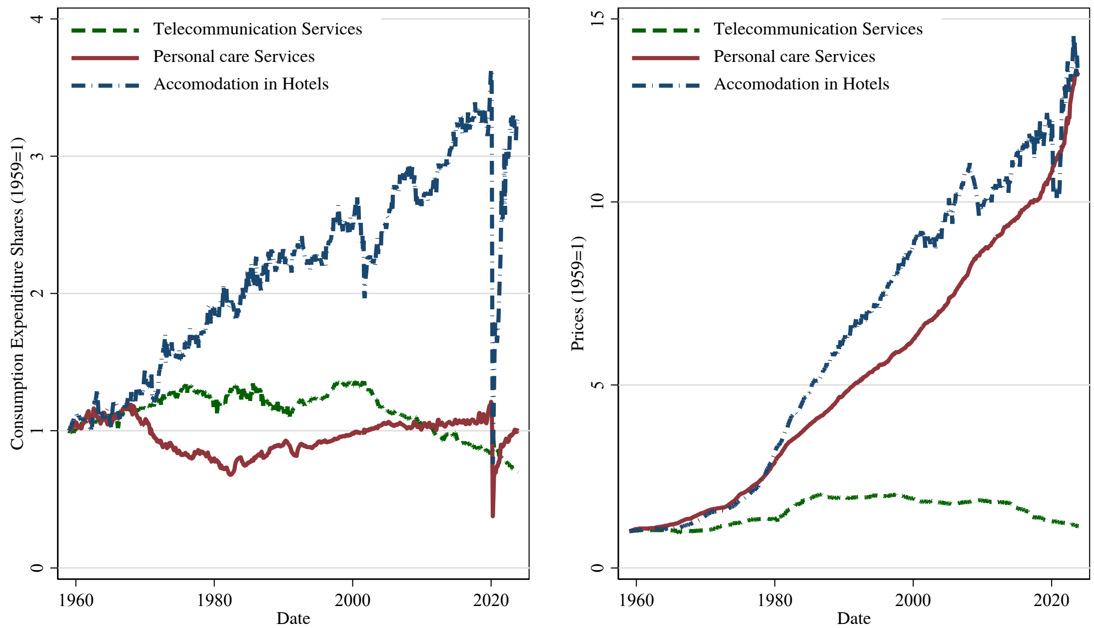


Figure 2: Expenditures and Prices of Some Services Categories Over Time. Source: Bureau of Economic Analysis.

'Automobile,' and 'Train Ticket,' for example, our consumer is acquiring a consumption outcome called 'Transportation Services.' These three products may also give rise to other consumption outcomes such as 'Conspicuous Consumption,' and 'Exercise.' The mapping of consumer products into consumption outcomes is the essence of the Lancaster (1966) framework. To it, we add what we call product *characteristics*. They are objects that govern the process through which the consumer converts purchased products into consumption outcomes. *Durability*, for example, is a *characteristic* of both a train ticket and a bicycle. The latter has more durability than the former, even though they both deliver the same consumption outcome. 'Brandedness' is another characteristic. It is likely to play a more important role for the bicycle and the automobile than for the train ticket. We describe this structure formally in Section 2, including what distinguishes our framework from that of Lancaster (1966).

We operationalize our framework by identifying a set of 23 different characteristics. We do so in the same manner as the O*NET, employing a mix of expert opinion, surveys, focus groups, and existing research on consumer behavior. These characteristics can, conceivably, be applied to describe *any* group of consumer products. We showcase our framework by using 69 products (including both goods and services) from the *Personal Consumption Expenditure* (PCE) data from the Bureau of Economic Analysis (BEA).[1] We assign numerical values to each product/characteristic pair. An automobile, for example, is assigned a high value for the characteristic 'Durability,'

[1]Our data are a partition of the aggregate consumption series published by the BEA; they add up to aggregate consumption. The BEA's most granular partition includes over 300 goods-and-services. We construct a 69-product partition for tractability reasons.

whereas a train ticket is assigned a low value. A wristwatch receives a high value for the characteristic 'Conspicuous Consumption,' whereas a gallon of gasoline receives a low value. The result of this procedure is that each product is defined in terms of a particular value of the 23-dimensional vector of characteristics that we define. If two products have identical numerical values for all 23 characteristics then, for our purposes, they are the same product.

The final ingredient of our framework is an organizational structure, a *taxonomy*, that houses our set of characteristics. We label our taxonomy as the *Ales-Lessem-Telmer-Zetlin-Jones (ALTZ) content model*. Its characteristics are organized into three categories, those that pertain to the consumer herself (*e.g.*, age, gender, geographic location), those that require information about preferences in order to operationalize, and those that do not.

The empirical section of our paper highlights as proof of concept three uses of the ALTZ content model. First, we show that the model helps organize the BEA's broad durable/non-durable/service taxonomy. Durability and storability emerge as central inputs into the BEA classification, but the analysis also highlights substantial heterogeneity within these coarse categories, especially within services. Second, returning to the behavior described in Figure 1, we use the ALTZ characteristics to explain cross-product expenditure movements during the COVID period. We show that the characteristic *Physical Human Interaction Required to Use or Consume*[2] is the most robust predictor of the collapse in expenditure shares during 2019-2020, while the recovery from 2020 to 2021 is additionally shaped by characteristics such as conspicuousness, tangibility, and rivalrousness. Third, we study longer-run expenditure reallocation between 1959 and 2023 within the service sector. There we show that the characteristic *Time-Investment Required for Use* helps explain which services have grown over time: services that require more time in use tend to gain expenditure share relative to those that do not.

Our paper is related to literature that treats the content of consumption, rather than the expenditure category alone, as central for behavior. Becker (1965); Lancaster (1966) emphasize that what matters for behavior are the time costs and characteristics embodied in consumption goods. Benhabib et al. (1991); Baxter and Jermann (1999) look at household production. They emphasize that behavior depends on the time and substitution margins embedded in them. Aguiar and Hurst (2005, 2007a) show that measured expenditure can move differently from consumption once shopping, home production, and time use are taken seriously. Closely related work in Gilboa et al. (2016); Hai et al. (2020) emphasizes that characteristics of goods such as memorability and salience matter to explain household consumption behavior. Taken together, these and related papers suggest that demand depends on a variety of characteristics (such as time intensity, search intensity, home-production content, and memorability) and not simply on whether an item is classified as a good or a service.

Buera and Kaboski (2012); Herrendorf et al. (2013); Boppart (2014); Comin et al. (2021) work

[2]A product requires physical human interaction for use if a person experiences close physical contact with at least one individual (outside their family) during the typical use of the product. Products that require physical human interaction may feature more (or less) intensive interactions, meaning that the interactions are with a greater (fewer) number of persons or for a longer (shorter) period of time.

on structural transformation models that emphasize combinations of nonhomothetic demand, relative price change, and systematic differences across sectors. Their common message is that expenditure shifts reflect stable differences in what categories require from households, including time, skill, and marketization, rather than a single homogeneous service margin. Relative to this literature, our contribution is to move to a systematic and explicit characteristic space analysis for consumption products.

The remainder of our paper is organized as follows. Section 2 provides our theoretical framework and the structure of our taxonomy. Section 3 provides descriptions of our set of characteristics and outlines how we measure them for each of our 69 BEA product categories. Section 4 provides a quantitative analysis of our set of characteristics, provides a validation of our approach, and provides an analysis of the U.S. consumption bundle during the time periods 2019-2021 and 1959-2023. Section 5 concludes.

2 Methodology

2.1 Products & Characteristics

Historically, taxonomies have been developed in several ways. One approach is to use data to derive the relevant characteristics; another is to specify those characteristics based on known underlying mechanisms (see Nickerson et al. (2013) for details). Hybrid approaches have also been adopted. We follow this latter approach. To explain our approach and set notation it is helpful to start with a standard consumption choice model. Consider an environment with M products.[3] Consumers have preferences $\tilde{U}$ defined over the set of products. Consumers are heterogeneous and are described by a vector of characteristics Ψ and an exogenous level of income Y. A consumer with income Y chooses products $\mathbf{q} \in \mathbb{R}^M$ taking prices $\mathbf{p} \in \mathbb{R}^M$ as given:

$$\max_{\mathbf{q}} \tilde{U}(\mathbf{q};\Psi), \quad s.t. \;\; \mathbf{p} \cdot \mathbf{q} \leq Y. \tag{1}$$

Note that the consumption space is very high dimensional. It represents every single product available in the product market. The consumer, therefore, is modeled as having a specific preference over every single element of this vast set.

Lancaster (1966) makes a compelling argument for an alternative approach, one aimed at both explaining the key drivers of consumption choice and at reducing the dimensionality of the consumer's problem. The main novelty of Lancaster's approach is the assumption that the utility function and the budget constraint are not defined on the same space. He re-writes the standard problem (1) as

$$\max_{\mathbf{c},\mathbf{q}} \; \widehat{U}(\mathbf{c};\Psi) \quad s.t. \quad \mathbf{c} = B \cdot \mathbf{q}, \qquad \mathbf{p} \cdot \mathbf{q} \leq Y \;\; , \tag{2}$$

where B is an $N \times M$ matrix and $\mathbf{c}$ is an N-dimensional vector of *consumption outcomes*.[4] The key

[3] Throughout, the word *products* is used as a stand-in for the more common phrase *goods and services*.

[4] Lancaster (1966) refers to what we call 'consumption outcomes' as 'characteristics.' In our framework the latter

difference is that $\widehat{U} : \mathbb{R}^N \to \mathbb{R}$ in (2), whereas $\tilde{U} : \mathbb{R}^M \to \mathbb{R}$ in (1), with $N < M$. Lancaster's consumers do not value products *per se*. Rather, they acquire products, $\mathbf{q}$, in the marketplace and then use them as inputs into a process that yields the consumption outcomes, $\mathbf{c}$, that they really care about. A meal in a restaurant, for example, would be an element of $\mathbf{q}$. Its monetary value appears in the consumer's budget constraint. Its consumption value, the utility it provides, arises from its ability to provide consumption outcomes such as 'nutrition,' 'aesthetic appeal,' and 'social interaction.' The matrix B describes how much the meal contributes to each of these consumption outcomes, along with how much other products contribute to the same set of outcomes. The product 'groceries,' for example, is distinct from the meal, has a different monetary value, but nevertheless contributes to the consumption outcome of 'nutrition.' This is the essence of Lancaster's reduction in dimensionality; many purchased products can contribute to the same consumption outcome.

Our goal is to build on the Lancaster framework and take it to the data. To do so requires more structure. We have two things in mind. First, Lancaster's $\mathbf{c}$ is, for the most part, unobservable. To observe the consumption outcomes generated by purchasing a restaurant meal (from the above example), one would need to measure things like 'social interaction' and 'aesthetic appeal.' Application of Lancaster's framework, therefore, either requires more structure on $\mathbf{c}$, or more structure on what gives rise to $\mathbf{c}$. In what follows, we attempt the latter. Second, there are a number of aspects of consumer products that are obviously important, but for which it is unclear how to incorporate them in the Lancaster framework. Take durability, for example. Is it a consumption outcome, $\mathbf{c}$, or is it something that governs the production of consumption outcomes? In the previous model, is it part of B or $\mathbf{q}$? With this in mind, we construct a consumer choice problem that both adds more structure and generalizes Lancaster's model (2).[5]

The distinguishing feature of our framework is the notion of product *characteristics*. Product characteristics, represented by the K-dimensional vector μ, are objects that govern how the consumer converts different bundles of purchased products $\mathbf{q}$ into consumption outcomes $\mathbf{c}$, so that $\mathbf{c} = f(\mathbf{q}; \boldsymbol{\mu})$ (where $\boldsymbol{\mu} = \{\mu_i\}_{i=1}^M$). Compared to Lancaster's framework, this might seem like no more than a change of notation. Nevertheless, we find it to be a useful organizational device in the following ways. First, it clarifies that our mapping from theory to data is inherent in assumptions about μ, not in assumptions about $\mathbf{c}$. As will become clear, this will be helpful in terms of observability. Second, our notion of characteristics gives rise to what is the crux of our paper, the development of an observable, O*NET-like taxonomy for consumer products. As the O*NET does for any occupation, we think of any consumer product as being a function of a relatively small number of characteristics. If products i and j have the same characteristics, so that $\mu_i = \mu_j$, then for our purposes they are the same product. Our taxonomy is simply an organizational

term is better used elsewhere, something that will become clear shortly.

[5]The goal of Lancaster (1966) and Lancaster (1971) is to establish an efficiency frontier for consumer products that allows empirical researchers to determine the substitutability across products *without* knowledge of consumer preferences. To establish this, his formulation requires strong modeling assumptions, including the requirement that products are linear combinations of consumption outcomes and that each consumption outcome be associated with a positive marginal effect on utility (see also Hendler (1975)). Our model does not feature such restrictions.

framework for the elements of μ.

We begin by partitioning μ into two sub-vectors, $\mu = \{\delta, \phi\}$. The vector δ contains what we call *objective characteristics*, objects that primarily influence consumption outcomes that are measurable without information about consumer preferences. *Durability* is an example of an element of δ. It is a characteristic of both a train ticket (low durability) and an automobile (high durability). It influences how these products give rise to the objectively-measurable consumption outcome 'transportation services.' Other examples of δ-characteristics include *Storability*, and *Degree of Human Interaction Required for Purchase*.

The second element of μ, ϕ, contains *subjective characteristics*. They influence consumption outcomes that are more subjective in nature, outcomes that are more likely to require information about preferences and that vary more across consumers. *Brandedness* is an example of an element of ϕ. The brand of a wristwatch, for example, influences a consumption outcome akin to the consumer's perception of their social status, something that is likely to require information on preferences in order to measure. Other examples of ϕ-characteristics include *Conspicuousness* and the extent to which *Keeping up with the Jones'* is an important aspect of consumer demand.

We can now be more precise about the definition of what a *characteristic* is in our framework. We define the set of consumption outcomes to include both objective and subjective outcomes, $\mathbf{c^o}$ and $\mathbf{c^s}$, respectively. The consumer's maximization problem is

$$\begin{aligned} &\max_{\mathbf{c}^o,\mathbf{c}^s,\mathbf{q}} \; U(\mathbf{c}^o, \mathbf{c}^s; \Psi) \quad s.t. \\ &\quad \mathbf{c}^o = f_I(\mathbf{q}; \delta), \quad \mathbf{c}^s = f_{II}(\mathbf{q}; \boldsymbol{\phi}), \quad \mathbf{p} \cdot \mathbf{q} \leq Y \;, \end{aligned} \tag{3}$$

where Ψ represents the set of observable characteristics that are specific to each individual consumer, such as age, gender and geographic location. The optimal solution to (3) can be written as $q(\mathbf{p}, Y, \boldsymbol{\mu}, \Psi)$. A *characteristic* is simply anything that affects the solution that is neither price, $\mathbf{p}$, nor income, Y.

Example 1. *To clarify our model, (3), we return to the transportation example from above. Bicycles, train tickets and automobiles are each elements of* $\mathbf{q}$*, the vector of quantities of purchased products. Each of the three products can contribute to multiple consumption outcomes, including both objective and subjective outcomes,* $\mathbf{c^o}$ *and* $\mathbf{c^s}$ *respectively. 'Transportation Services' is one element of the former. Each product provides the consumer with transportation services, but in different ways that are captured by the characteristic,* δ*. 'Durability' is one such characteristic. It captures the fact that the bicycle and the automobile provide a stream of transportation services, whereas the train ticket does not. This process is captured by the function* f_I*, with* $\mathbf{c}^o = f_I(\mathbf{q}; \delta)$ *mapping elements of* $\mathbf{q}$ *into elements of* $\mathbf{c^o}$*, governed by* δ*. The three products can also contribute to other consumption outcomes. For example, a subjective outcome is 'Conspicuous Consumption,' an element of* $\mathbf{c^s}$*. Certain types of automobiles provide more conspicuous consumption than does a bicycle or a train ticket. This process is governed by the function* $f_{II}(\mathbf{q}; \boldsymbol{\phi})$*, where 'Conspicuousness' is an element of the vector of subjective characteristics,* ϕ *for each good. This example illustrates the sense in which certain characteristics of products are more directly connected than others to* *the consumption outcomes that the consumer ultimately derives utility from.*

Our model provides a low-dimensional representation of the product space; products are described by the low-dimensional vector of characteristics they possess. This leads to an organizational device–a taxonomy–with which we can house the set of characteristics that constitute a basis for the product space.[6] We expand upon this next.

2.2 A Taxonomy for Products

Following the O*NET, we refer to our taxonomy as a *Content Model* and label it as the Ales, Lessem, Telmer, Zetlin-Jones (2025) (ALTZ) *Content Model*. It is depicted in Figure 3. Its skeleton is defined along three vertices, each of which represents one element of the universe of characteristics, δ, ϕ and Ψ, from our model (3). At the top of the diagram are objective characteristics, δ, those that are most closely related to products themselves. At the bottom are characteristics related to consumers and to consumer preferences. On the left-hand side are characteristics of the consumer herself, Ψ, and on the right-hand side are subjective characteristics, ϕ, those that are more closely related to consumer preferences. We refer to each of these three sets of characteristics as a *domain* of our content model.

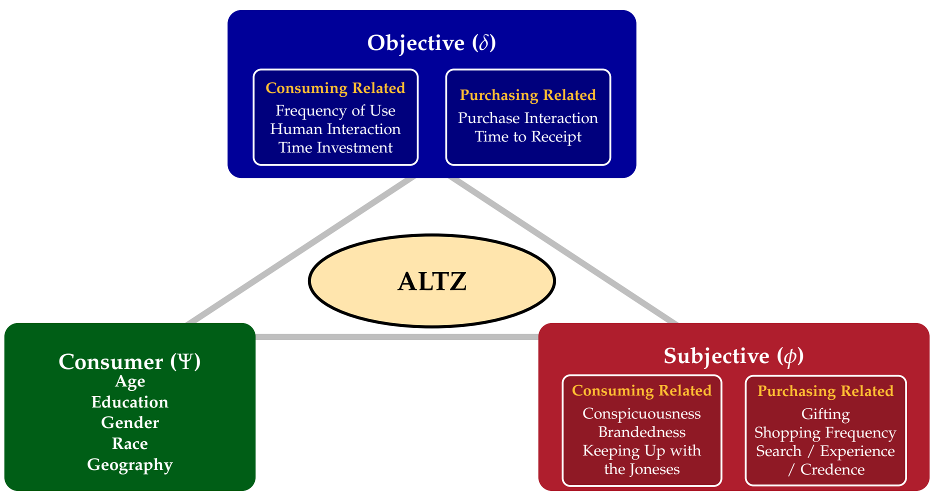


Figure 3: The Ales, Lessem, Telmer, Zetlin-Jones (2025) (ALTZ) Content Model including a representative subset of characteristics in each category.

Figure 3 lists a subset of the characteristics that populate each of our model's three domains. Table 1 in Section 3 provides the entire set. Next, we elaborate on each of the domains in Figure 3 and provide some representative examples. Note that, while some characteristics align closely with a specific domain (vertex), others more likely lie between two of the three vertices. We begin with the top-most vertex and proceed in a counter-clockwise direction.

[6] Lancaster (1966) highlights how the standard model is not a useful starting point for building a product taxonomy as each product is different only with respect to preferences. While Lancaster (1966) recognizes the possibility of using (2) to build a taxonomy, he does not operationalize this step.

- *Objective Characteristics (δ).* These characteristics are primarily related to the products themselves, as opposed to consumer preferences. They modify the process described by the function f_I, which converts bundles of products into *objective* consumption outcomes Examples of objective characteristics include *Durability*, *Excludability*, and *Rivalrousness*. We find it useful to distinguish objective characteristics according to whether they are primarily associated with consuming or purchasing the product in question.

  Subdomains of δ

  - *Consuming Related.* Consider the COVID pandemic, for example. Using an automobile posed a smaller risk of infection than using a train ticket. The same was true of a restaurant meal as compared to groceries. The characteristic *Physical Human Interaction Required to Use/Consume* therefore modifies the process of converting these products into objective consumption outcomes according to the likelihood of contracting the COVID virus.
  - *Purchasing Related.* Continuing the previous example, we include in this category characteristics such as *Physical Human Interaction Required to Purchase/Receive*. *Time Required Between Purchase and Receipt* is another characteristic that is clearly in this category.

- *Characteristics of the Consumer (Ψ).* Consumer demand, $q(\mathbf{p}, Y; \phi, \Psi, \delta)$, depends directly on characteristics of the consumer herself: age, education, gender, race, geographic location and so on.

- *Subjective Characteristics (ϕ).* These characteristics are manifest in the process of converting products into consumption outcomes that are more closely related to consumer preferences. In this sense, these characteristics are more closely aligned with the consumer herself rather than with the specific product. As with objective characteristics, it is useful to distinguish these characteristics according to whether they are more or less associated with consuming versus purchasing.

  Subdomains of ϕ

  - *Consuming Related.* Examples include *Conspicuousness, Brandedness*, and *Keeping up with the Joneses*. Returning to the automobile example, a consumer may derive utility from the car itself, rather than the set of objective consumption outcomes that a car delivers. A luxury car, for example, may yield (subjective) utility in the form of conspicuous consumption. The same is true of products that reflect a large degree of *Brandedness*.
  - *Purchasing Related.* One example is 'Gifting Motive.' Some products, flowers for example, are purchased in order to generate the subjective consumption outcome of the warm glow of gift giving. A second example is the characteristic 'Search, Experience, Credence,' which measures the means with which a consumer can ascertain product features that are related to quality.

3 Measurement

To implement our content model in practice, we proceed in three steps: *(i)* define the set of characteristics $\{\phi, \delta, \Psi\}$; *(ii)* define the set of products to be represented in terms of the characteristics; *(iii)* obtain quantitative estimates of the characteristics for each of the products.

3.1 Defining the Set of Characteristics

For the O*NET, the development of the characteristics for various occupations evolved organically over many decades.[7] To generate the list of characteristics and their definitions, we combine prior knowledge, expert analysis, and survey data. In implementing this step, we define the scope of the taxonomy so that it is: *(i)* parsimonious, avoiding unnecessary or overlapping characteristics; *(ii)* comprehensive, with characteristics that are not tied to any specific research question; and *(iii)* explanatory, meaning that the characteristics should exhibit meaningful variation across different goods.[8] In practice we proceed as follows:

1. As a starting point, and echoing the expert analysis approach of the O*NET, we used our own expertise to include and define certain characteristics.

2. Whenever possible we relied on existing literature discussing, in isolation, product characteristics. For example, we considered the literature on 'memory goods' (e.g., Gilboa et al. (2016), Hai et al. (2020)) or the framework based on search, experience and credence goods (e.g., Ford et al. (1988)).

3. Finally, we assembled a focus group of Carnegie Mellon University undergraduates. The students, supervised by us, formulated some of the elements of our characteristic set based on both personal experience and their own research. In addition, the students surveyed their immediate networks to determine any gaps in the set of characteristics.

This process resulted in a final set of 23 characteristics, not including demographic characteristics that we directly associate with the consumer herself. The characteristics are listed in Table 1.

Our companion technical document provides granular details for how each characteristic is defined and measured.[9] Overall, we view our characteristics in much the same way as the profession has come to view the O*NET's occupational characteristics, as an evolving dataset that evolves based on expert analysis and empirical validation.

[7] The precursor of the O*NET, the *Dictionary of Occupational Titles* (DOT), was initiated in 1939. By the 1990s the number of titles in the DOT grew to roughly 12,000. In 1996 the O*NET was born as the offspring of the DOT. The number of occupational titles was condensed to a list of 1,102, which was linked to a taxonomy of occupational characteristics that now numbers 277.

[8] For further discussion of approaches to building taxonomies, see Bailey (1994).

[9] The technical document is available here: CPCPTechnicalDocument.pdf.

Table 1: Twenty-Three Characteristics of the ALTZ Content Model.

Domain and Characteristic	Abbreviation
Objective Characteristics	
Tangibility	Tangible
Excludability	Exclude
Rivalrousness	Rival
Durability	Durable
Storability	Storable
Within-Product Heterogeneity	ProdHet
Predictable Time Variation: Production	TVProdcn
Consuming Related	
Frequency of Consumption/Use	ConsFreq
Physical Human Interaction Required to Use or Consume	ConsHuman
Time-Investment Required for Use	ConsInv
Purchasing Related	
Physical Human Interaction Required to Purchase or Receive	PurchHuman
Time Required Between Purchase and Receipt	PurchTimeLag
Subjective Characteristics	
Consuming Related	
Conspicuousness	Conspic
Keeping Up With the Joneses	KUWtJ
Brandedness	Brand
Subsistence Level	Subsist
Susceptibility to Habit Formation	Habit
Predictable Time Variation: Preferences	TVPrefs
Purchasing Related	
Gifting Products	Gift
Philanthropy and the Provision of Public Goods	Charity
Shopping Frequency	ShopFreq
Time-Investment Required for Purchase	PurchInv
Search, Experience, Credence	SEC

3.2 Defining the Set of Products

Our set of products derives from BEA Table 2.4.5U, *Personal Consumption Expenditures by Type of Product*. The goal is to settle on broad consumption categories that balance disaggregation and parsimony. The traditional BEA taxonomy of three categories - durables, non-durables, and services - leans too far toward parsimony. The highest level of disaggregation in the BEA table contains 366 goods and services, leaning too far in the other direction. We settled on 69 distinct products and aggregated the remaining categories accordingly. The final list of products is in Appendix C.

3.3 Estimating Characteristics

We follow several different approaches to associate numerical values with each of our 23 characteristics, for each of the 69 products described above. For *Time-Investment Required for Purchase* and *Shopping Frequency* we used the consumption surveys in the *Consumer Expenditure Survey*. For *Durability* we began with the traditional BEA taxonomy and then used our own judgment to expand the scope to include, for instance, some BEA-defined services such as health care, which are obviously durable-good expenditures. Finally, for the remaining characteristics, we adopted the expert-assessment approach used by O*NET, with the authors serving as the expert evaluators. For each characteristic, we specified, *a priori*, an algorithm for assigning values across products. We then implemented these algorithms through multiple iterative rounds, incorporating discussion and feedback at each stage to refine the estimates. Additional details are available in our companion technical document: CPCPTechnicalDocument.pdf.

The values we assign are integers between one and five, with the scale for each product-characteristic pair depending on the characteristic (all data are normalized prior to analysis). Smaller values indicate that the intensity of a particular characteristic for a particular product is small. Take for example the characteristic *Conspicuousness*. The product *Gasoline, lubricants, fuel, and other energy products* is unlikely to be associated with conspicuous consumption. Therefore it is assigned a value of one. The product *Jewelry and Watches*, on the other hand, is a common example of a conspicuous-consumption product. It is assigned a value of five. A different characteristic results in different values for these products. For the characteristic *Shopping Frequency* gasoline etc. gets a high value and jewelry gets a low value. Overall, this procedure results in a 69×23 matrix of numerical values, where the 69 rows represent products and the 23 columns represent characteristics.

Example 2. *We now return to our previous example of how automobiles and public transportation are consumer products that are similar along some dimensions and different along others. A similarity is that they both provide transportation services. However, this is* not *what our framework is designed to measure. Transportation services are an example of a consumption outcome,* $\mathbf{c}^{\circ}$, *from Equation (3). They are part of the 'bundle' of consumption objectives that one acquires when purchasing a car. Measuring every item of every product-specific bundle is not feasible. What we are able to measure are* characteristics, *a lower-*

dimensional, observable set of variables that affect consumer demand for each product. An important one for this example is Durability*. It is assigned a high value for* Automobiles *and a low value for* Public Transport*. These products are very different along the durability dimension. The same is true for the characteristic* Keeping Up With the Joneses*. In contrast, cars and train rides are assigned similar values for the characteristic* Search, Experience, Credence*, which is a measure of how a consumer assesses product quality. The same is true for the characteristic* Subsistence*.*

3.4 Relationship Across Characteristics

With the estimated characteristics we can now have a closer look at the relationship across them. To do so we use the information contained in the content model that evaluates our 23 characteristics over the 69 aggregated BEA products. Before proceeding we standardize the values for each characteristic. Overall, we observe cross-correlations across characteristics ranging from -0.54 to 0.76. While we observe no perfect correlation across characteristics, certain characteristics are closely related.[10] For example, not surprisingly, *Durability* and *Storability* are the most correlated characteristics while the most negatively correlated characteristics are (also not surprisingly) *Time-Investment Required for Purchase* and *Shopping Frequency*. Figure 9 in Appendix A displays the entire correlation matrix.

3.5 Validation

As in O*NET, our approach to measuring characteristics relies on expert opinion. To provide additional validation of our approach, we partnered with a professional data firm *Civic Science* (CS) to compare the value of our characteristics with ones determined by a survey. CS manages a network of web-based polling applications distributed across third-party websites to engage consumers in attitudinal research.[11]

For budgetary reasons we limit our survey to three characteristics (*Conspicuousness, Physical Interaction While Purchasing/Receiving, and Physical Interaction While Consuming/Using*) and eighteen products. The list of products surveyed is in Appendix B. To illustrate the data collected by CS for each product-characteristic pair, we present an example of the survey question and the resulting data for the characteristic *Conspicuousness*.[12]

> **Conspicuousness - Survey Question**: *How much do people consider how purchasing* <product name> *will influence other people's opinions of them?*

The responses from just under 3,000 respondents for two of the eighteen products surveyed are displayed in Table 2.

[10] Of course the degree of correlation depends critically on the number and type of products. It is clear that working with a larger number of more detailed product categories would lower the correlation between characteristics while working at a more aggregated level would raise the correlation.

[11] For additional detail on the methodology adopted by *Civic Science* refer to the following whitepaper Link.

[12] The formal definition of *Conspicuousness* from our technical document is: *A conspicuous consumption product is one*

Table 2: Survey on **Conspicuousness** - Percentage of Respondents for Each Alternative.

	Nursing Home Care	**New and Used Cars and Trucks**
Very much	12.7	18.0
Somewhat	21.6	27.3
Very little	23.8	22.1
Not at all	26.9	20.6
I'm not sure	15.1	11.9
Total responses	2,950	2,966

Note that the survey question used is much simpler than the complete definition of the characteristics as described in our technical document. The rationale behind this is to ensure that survey respondents have a clear grasp of what is being asked. The development of the survey instrument was a collaborative effort with the survey expert at CS.

Figure 4 compares the values of the characteristics from the ALTZ content model with the averages from the survey. A positive correlation is observed for all three characteristics. The relationship is stronger for the characteristic *Physical Human Interaction Required to Purchase or Receive* (*PurchHuman*) with a correlation of 0.71 and weakest for *Physical Human Interaction Required to Use or Consume* (*ConsHuman*) with a correlation of 0.3.

4 Applying the ALTZ Content Model

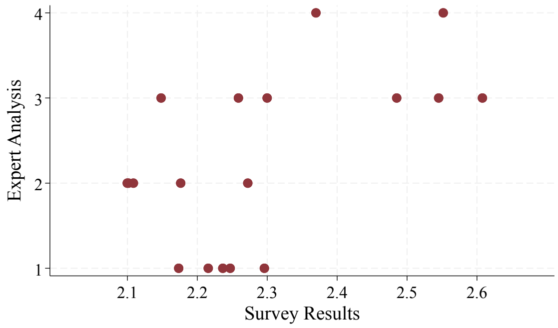
In this section, we highlight several use cases for the ALTZ content model. The first application involves using the estimated characteristics to connect them with established classifications such as the BEA consumption expenditure categories.[13] The second application focuses on explaining patterns in consumer spending over the pandemic recession. The third application takes a long-run perspective looking at how the consumption basket has changed over time.

4.1 Explaining Durables, Non-Durables and Services

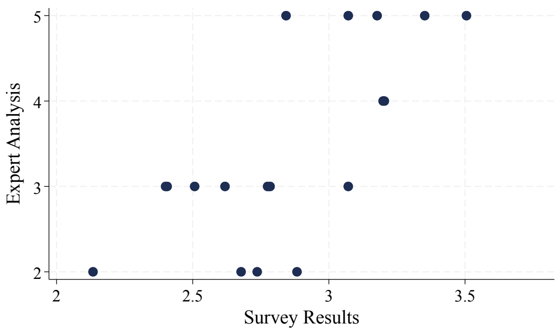
The standard BEA classification usually considers products divided into one of three broad categories: durables, non-durables, and services. The definition from the BEA states that durables are: *"Tangible products that can be stored or inventoried and that have an average life of at least three years."* The BEA definition, from the perspective of the ALTZ model, overlaps several characteristics of products that our content model identifies. For example, the definition includes criteria based on *Tangibility*, *Storability*, and *Durability*. Similarly, services are defined by the BEA as *"Products that cannot be stored and are consumed at the place and time of their purchase;"* this definition overlaps with product characteristics such as *Storability* and *Time Required Between Purchase and Receipt*.

for which the consumer perceives that owning/using the product will have an important effect on other people's perceptions of the consumer. Perceptions can be those of admiration or those of stigma.

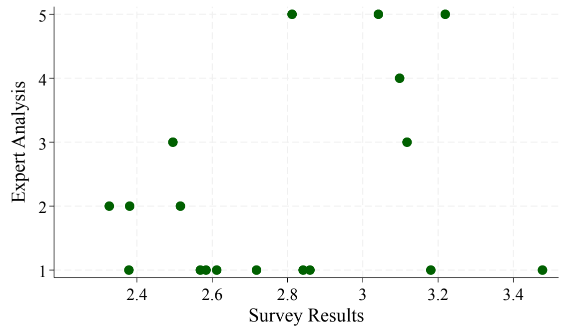
[13] In Appendix A we also look at how the content model provides a rationale for consumer behavior by relating characteristics to estimated price and income elasticities across goods and services.



(a) Characteristic: *Conspic*



(b) Characteristic: *PurchHuman*



(c) Characteristic: *ConsHuman*

Figure 4: Scatter Plot of ALTZ Content Model Values from Expert Analysis versus Survey Results. Figure **(a)** plots values for Conspicuousness (*Conspic*), Figure **(b)** plots values for Physical Human Interaction Required to Purchase or Receive (*PurchHuman*), and Figure **(c)** plots values for Physical Human Interaction Required to Use or Consume (*ConsHuman*).

We can use the characteristics in the taxonomy to better understand the classification among durables, non-durables and services of the BEA. We look at the 69 aggregate products.[14] A difficulty in expressing the relationship is that each product is described by a point in a 23-dimensional space (each dimension representing each characteristic). To aid in the visualization we use a t-distributed stochastic neighbor embedding algorithm (t-SNE) to reduce the dimensionality of the dataset to two dimensions. The results are displayed in Figure 5; in the diagram products close to each other are "close"in the 23-dimensional space. Each product is labeled with a number signifying the ALTZ product number (refer to Appendix C for a description of each number). In the figure we also overlay the BEA classification with different colors. While information on BEA classification is not used in the algorithm, the t-SNE plot seems to generate clusters of similar products. For example, *durables* appear close to each other. The diagram highlights the high heterogeneity of *services* (this heterogeneity is something we will explore further in Section 4.3). Indeed, some *services* appear to be more closely related to *durables* than to other types of *services*. Compare, for example, *Accommodations in hotels and motels* (product number 49)

[14]Refer to the online appendix for a detailed crosswalk between BEA products and the aggregated products used in this project.

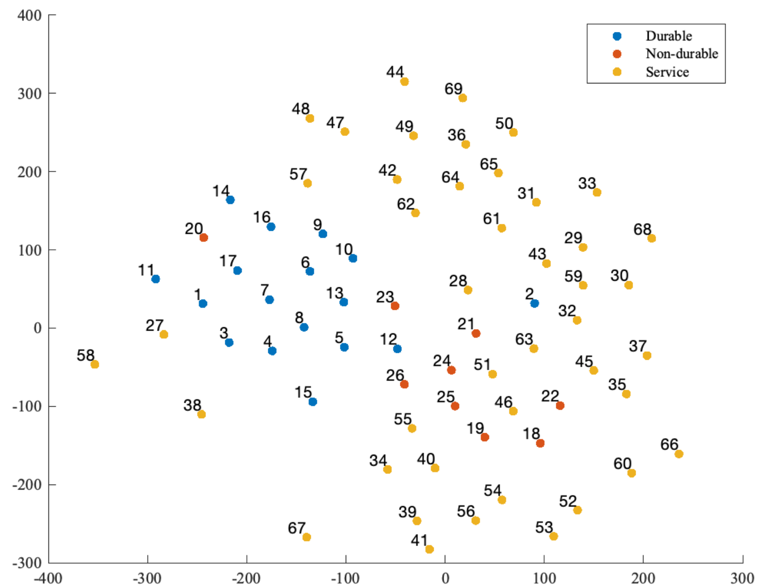


Figure 5: t-distributed stochastic neighbor embedding of ALTZ Products and Characteristics. Numbers in the figure represent the ALTZ product numbers.

and *Insurance* (product 53).[15] To formally relate our characteristics to the ones used by the BEA, we use a decision-tree, random forest classifier (Breiman (2001)). The idea is that in a decision tree we can use our characteristics to classify whether a product is a durable, non-durable, or service based on the BEA definition. Ideally only a few of the characteristics should be sufficient to classify products. The random forest algorithm provides a measure of importance of each variable (in our case the characteristics) when trying to build a classifier for the type of BEA products. The results are displayed in Figure 6 (refer to Table 1 for reference on the abbreviations used in the Figure). We see here that the algorithm recognizes *Durability* and *Storability* as key variables in a decision tree to determine whether a product is a durable, non-durable, or service based on the BEA definition. Indeed above we highlighted how these two characteristics are implicit in the definition of types of products used by the BEA. However, from Figure 6, we also see that many characteristics are needed for a correct classification of products. This points to the large heterogeneity in products in each of the three classifications. This result also serves as a warning in analyzing consumption expenditures, especially services, using only the BEA classification.

[15]Care should be used in interpreting t-SNE plots as they try to represent high dimensional objects in lower dimension. As an example of potential distortion induced in this process, consider that the closest products (in terms of Euclidean distance in the original 23-dimensional space) are *Video and audio streaming and rental* (41) and *Television services* (39); this seems to be consistent with the diagram in Figure 5. On the other hand, the furthest apart products are *Educational services, elementary and secondary schools* (58) and *Gasoline, lubricants, fuel and other energy products* (21); these are not the furthest apart products in Figure 5.

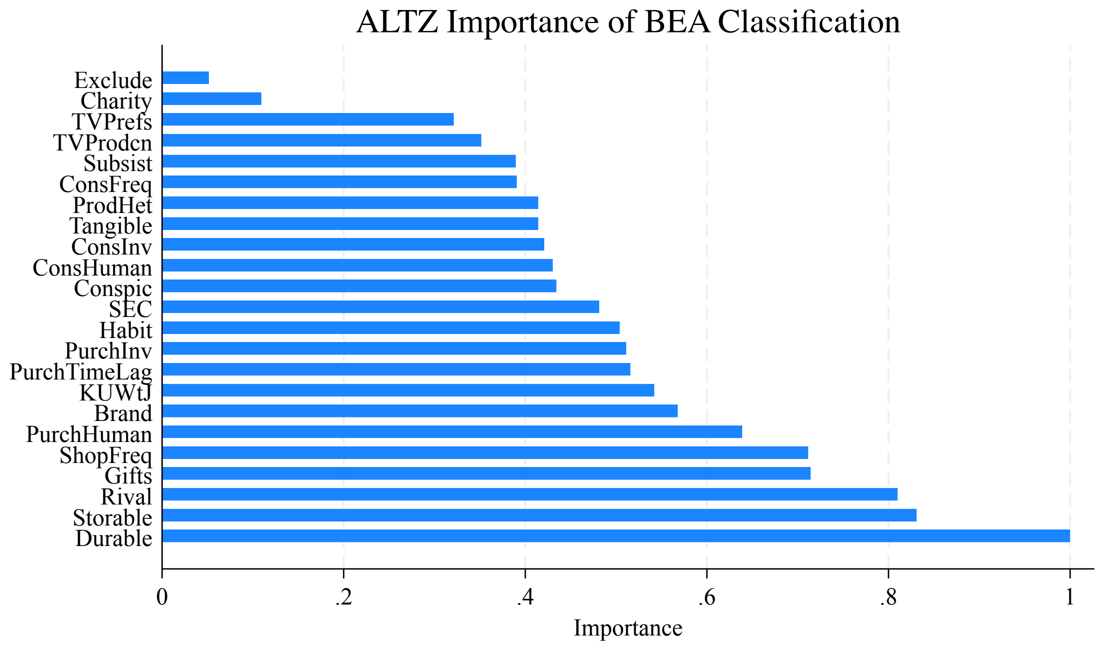


Figure 6: Random Forest Importance of Characteristics in Determining BEA Product Classification.

4.2 Aggregate Consumption Analysis: The 2019-2021 Pandemic

The framework developed in this paper can help us understand the evolution of consumption patterns over time. We begin by looking at the behavior of aggregate consumption during the 2019-2021 COVID pandemic. As anticipated in the introduction, the 2019-2021 period saw a large collapse and quick recovery of durables and, differently than previous recessions, a large collapse and slow recovery for services. The empirical exercise proceeds in two steps. We first recover product-level preference shifters from observed movements in expenditure shares and prices. We then ask whether the cross-section of these preference shifts can be explained by the product characteristics in the ALTZ content model. Our starting point is the model described in (3) substituting out the consumption outcomes from the utility function. In what follows q_{it} is real expenditures on the i-th product at time t. Our parametrization of the aggregate consumer problem at time t is given by:

$$\max_{q_{1t},\ldots,q_{Mt}} \left[\sum_{i=1}^{M} a_{it}^{\frac{1}{\sigma}} q_{it}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

subject to:

$$\sum_{i=1}^{M} p_{it} q_{it} = Y_t, \forall t.$$

Here, p_{it} is the price for consumption good i in period t and Y_t denotes financial resources available in period t. Relative to the original problem in (3) we now allow the utility function U

and the function governing the relationship between goods and consumption outcomes to vary over time. In a simple way, in the objective function, the weights a_{it} now represent the preference shifter for good i at time t. We next show how to recover weights a_{it} from aggregate consumption data. From first order conditions of the above problem we get:

$$s_{it} = \frac{a_{it} p_{it}^{1-\sigma}}{\sum_j a_{jt} p_{jt}^{1-\sigma}}, \quad \forall i = 1, \ldots, M, \forall t. \tag{4}$$

Here, s_{it} is the share of total expenditures incurred in good i. Taking logs and differentiating between two time periods t and t' we get:

$$\log \frac{s_{it'}}{s_{it}} = \log \frac{a_{it'}}{a_{it}} + (1-\sigma) \log \frac{p_{it'}}{p_{it}} + \Gamma_{t',t}, \quad \forall i = 1, \ldots, M. \tag{5}$$

Here, $\Gamma_{t',t}$ is a constant depending on t and t'. The above equation can be used to recover the relative weights for consumption over time. As a final step we connect the recovered preference shifters to our 23 characteristics by estimating the following:

$$\log \frac{a_{it'}}{a_{it}} + \Gamma_{t',t} = \beta_0 + \sum_{j \in J} \beta_j^c \cdot X_{i,j}, \quad \forall\ i = 1, \ldots, M. \tag{6}$$

In the equation above, $X_{i,j}$ denotes the value of characteristic j for product i in the content-model data. In (6) the coefficients β_j^c denote the role that the $j-$th characteristic has in explaining movement in the preference shifters.

To study the 2019-2021 time period, we are interested in both the initial drop and the subsequent recovery. For the "drop" we set t' as April 2020 and t as April 2019. For the "recovery" we set t' as April 2021 and t as April 2020. Our data source for expenditures and prices is the BEA.[16] Finally, we use a value of $\sigma = 0.89$ taken from Herrendorf et al. (2013) (in Appendix A we consider the Cobb-Douglas case with $\sigma = 1$). To start, Table 3 and Table 4 look at an estimation of equation (6) on fewer chosen characteristics. The selection of the first set of characteristics is motivated by the different behavior of durables, non-durables and services during the pandemic. The characteristics, given the results in Figure 6, represent the three most important characteristics in explaining the BEA product classification. The remaining two characteristics, motivated by the nature of the pandemic itself, relate to the interaction with other humans during the consumption (*ConsHuman*) and purchase (*PurchHuman*) of products.

Table 3 summarizes the beginning of the pandemic, the "drop." In column (1) of Table 3 we see how the top characteristics used to explain the BEA product classification fail at providing any significant explanatory power in the behavior of products during the pandemic. In column (2) we add the characteristics relating to interaction with humans. In this case the characteristic *ConsHuman* is a highly significant variable (notably interaction in purchase, *PurchHuman* is not

[16] Specifically we use table 2.4.4U for Price Indexes and Table 2.4.5U for Personal Consumption Expenditures by Type of Product. We look at data reported at a monthly frequency.

Table 3: Estimates of β_j^c in equation (6) between April 2019 and April 2020.

	(1)		(2)	
Durable	0.187	(0.111)	0.209*	(0.0909)
Storable	0.0623	(0.110)	-0.0409	(0.0901)
Rival	-0.0975	(0.0750)	-0.0834	(0.0597)
ConsHuman			-0.293***	(0.0836)
PurchHuman			-0.0965	(0.0840)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

significant). The estimates for *ConsHuman* are intuitive: a higher degree of human interaction in consumption is associated with a decrease in expenditure shares during the drop phase of the pandemic. This result is also reminiscent of Cirelli and Gertler (2025) when splitting production in the *contact* and no-contact sector for the economy.

Table 4: Estimates of β_j^c in equation (6) between April 2020 and April 2021.

	(1)		(2)	
Durable	-0.0919	(0.0770)	-0.0901	(0.0656)
Storable	-0.0215	(0.0762)	0.0330	(0.0650)
Rival	0.162**	(0.0518)	0.158***	(0.0431)
ConsHuman			0.235***	(0.0603)
PurchHuman			0.00182	(0.0606)

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4 looks at later stages in the pandemic, the "recovery." As in the previous case, *ConsHuman*, together with Rivalrousness (*Rival*), plays an important role in explaining how different products recovered after the initial set of pandemic lockdowns.

The hand-picked specifications are useful for testing hypotheses motivated directly by the nature of the pandemic. A natural next question, however, is whether the same qualitative conclusions emerge when the data are allowed to choose among the full set of 23 characteristics. To complement the hand-picked specifications, we estimate equation (6) with a linear lasso using the full set of 23 characteristics. We use two penalization rules for each outcome: an adaptive lasso and a plugin lasso with a more conservative penalty.[17] The tables report postselection coefficients, so blank cells indicate characteristics that are not selected, and the bottom rows summarize

[17] Here "adaptive lasso" refers to the specification selected by cross-validation with data-dependent penalty weights, so regressors with stronger preliminary signals are penalized less heavily. "Plugin lasso" refers to the specification in which the penalty is chosen using a theory-based plug-in rule rather than cross-validation; in our application it is more conservative and therefore tends to select fewer characteristics.

the number of selected characteristics and the selected value of λ. For the pandemic drop, *ConsHuman* is the only characteristic retained by both procedures, making the degree of face-to-face interaction in consumption the most robust predictor of the collapse in expenditure shares. The adaptive lasso also selects *Conspic*, *Subsist*, *Charity*, *Tangible*, *Durable*, and *ConsFreq*, but these do not survive the stricter plugin penalty. For the recovery, *Conspic* is selected by both procedures, the plugin lasso additionally keeps *Tangible*, and the adaptive lasso selects a broader set that also includes *Rival*, *Durable*, *ConsHuman*, and *ConsInv*. Taken together, the lasso results suggest that the downturn was most tightly linked to human-contact intensity in consumption, whereas the rebound loads more on conspicuousness and broader product-type traits.

Table 5: Lasso estimates of β_j^c in equation (6) for April 2019 to April 2020. Columns report adaptive lasso and plugin lasso.

	Adaptive lasso	Plugin lasso
Conspic	-0.164	
Subsist	0.087	
Charity	0.148	
Tangible	-0.100	
Durable	0.233	
ConsFreq	0.150	
ConsHuman	-0.207	-0.386
Constant	-0.211	-0.211
Selected characteristics	7	1
λ	0.0779	0.4348

Table 6 reports the recovery-period estimates. A useful way to read the two lasso tables together is that the collapse and the recovery are not mirror images of each other. During the initial drop, the most robust signal is the negative role of human interaction in use, consistent with lockdowns and social-distancing constraints. During the recovery, more product-specific margins reappear: conspicuousness, tangibility, and rivalrousness matter more, suggesting that the normalization of social and in-person activity brought back demand especially for goods and services whose value depends on visibility, physical presence, or limited sharing. This asymmetry is consistent with the idea that the pandemic first shut down contact-intensive consumption broadly, and only later allowed demand to reallocate along finer dimensions of product content. In Appendix A, we consider the Cobb-Douglas case ($\sigma = 1$); similar results hold in this case as well.

Table 6: Lasso estimates of β_j^c in equation (6) for the pandemic recovery, April 2020 to April 2021. Columns report adaptive lasso and plugin lasso.

	Adaptive lasso	Plugin lasso
Conspic	0.094	0.183
Subsist	-0.098	
Charity	-0.059	
Tangible		0.132
Rival	0.149	
Durable	-0.114	
ConsHuman	0.154	
ConsInv	0.069	
Constant	0.160	0.160
Selected characteristics	7	2
λ	0.0215	0.4348

4.3 Aggregate Consumption Analysis: Historical Trends For Services

We next take a longer time-horizon perspective. As in the previous section, we look at BEA consumption and price data but now consider the time period between 1959 and 2023. Similar to the previous section we are interested in extracting information on preference shifts as described in equation (5) and project this information on characteristics as described in equation (6). In equation (5) we now consider t as the average of the time period 1959-1964 and t' as the average during 2018-2023.

In Figure 7 we display the changes over time of prices and consumption shares. In the figure we observe well-known trends on the evolution of product categories over time: the general decline of durables and non-durables and the shift of the consumption basket toward services (see for example Herrendorf et al. (2013)). Looking more closely we also see a wide dispersion in the growth behavior of services. This is reminiscent to what we saw earlier when looking at the wide dispersion of services in our characteristic space: classifying a product as a service is broader than the classification with respect to durables and non-durables and thus less informative. We use our content model to provide additional insights on why certain services have grown over time and some have not.

Similar to what we have done in the previous section, we look at the relationship between the log $\frac{a_{it'}}{a_{it}}$ and the 23 characteristics. We focus on the BEA-defined service sample. Lasso is then estimated without a constant for two measures of the preference shift: one fixing $\sigma = 0.89$ and one recomputing log $\frac{a_{it'}}{a_{it}}$ with our own estimated elasticity value of $\sigma = 0.96$. Across all four lasso specifications, *Time-Investment Required for Use* (*ConsInv*) enters positively and is the only characteristic selected by both plugin lasso columns, making it the most robust predictor of long-

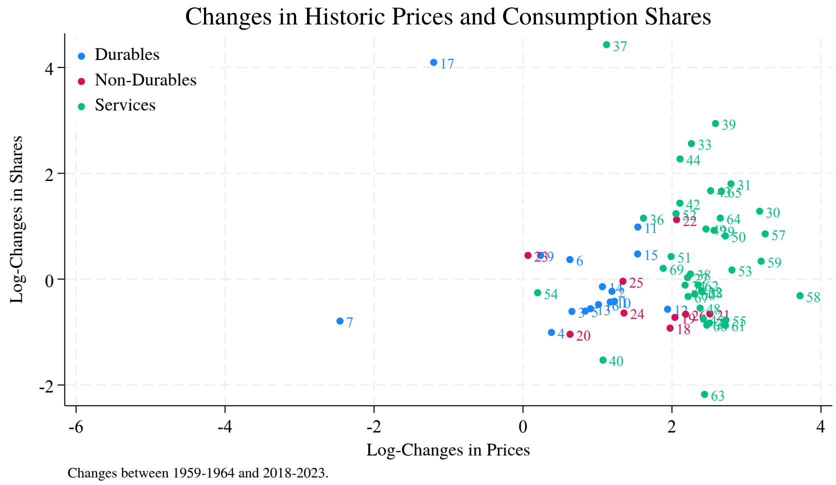


Figure 7: Historical changes in Prices and Consumption Shares. Numbers represent Product numbers for our content model.

run growth in service demand. The adaptive lasso additionally selects *SEC* and *ConsFreq*, both with negative coefficients, but these disappear under the plugin penalty and therefore appear less robust than the positive role of time investment in use. Restricting attention to BEA-defined services is important for interpretation. Once we condition on being a service, the characteristics that separate services from durables and non-durables become less informative, and the interesting variation is instead within the service category itself. The lasso results show that this within-service heterogeneity is economically meaningful: long-run growth is better organized by differences in how services are consumed, especially by how much time they require, than by the coarse BEA service label alone.

Table 7: Lasso estimates of β_j^c in equation (6) for BEA-defined services between 1959–1964 and 2018–2023. Columns (1)–(2) fix $\sigma = 0.89$ and columns (3)–(4) use the estimated $\sigma = 0.96$; within each pair we report adaptive and plugin lasso results.

	Fixed $\sigma = 0.89$		Estimated σ	
Characteristic	Adaptive	Plugin	Adaptive	Plugin
SEC	-0.442		-0.419	
ConsFreq	-0.453		-0.521	
ConsInv	0.556	0.654	0.585	0.710
Selected characteristics	3	1	3	1
λ	0.4207	0.5782	0.0935	0.5782

In this exercise, *Time-Investment Required for Use* (*ConsInv*) emerges as significant. That is, products that tend to score high in this characteristic, perhaps not surprisingly, see their share of consumption grow over time. For a visual representation of this finding, Figure 8 replicates Figure 7 focusing on services and coloring the different products relative to their score for *ConsInv*. The relationship is visible here too, especially when comparing products that score the least in terms of requiring time investment (*ConsInv*=1) to those that require the most (*ConsInv*=3).

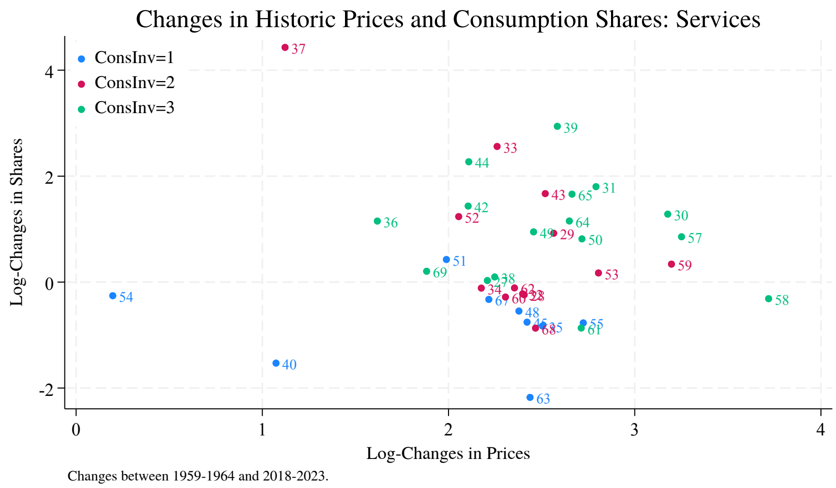


Figure 8: Historical changes in Prices and Consumption Shares for Services. Numbers represent Product numbers for our content model.

The findings highlight how the time available for consumption matters. Indeed, products that require more time are those that have grown the most over time. This is intuitive given the historical move in the U.S. away from work hours and toward education and leisure (see for example Aguiar and Hurst (2007b) and Ramey and Francis (2009), for documentation of this finding).

5 Conclusion

The paper introduces a comprehensive framework for characterizing consumer goods and services through a reduced set of 23 characteristics, offering a refined lens for analyzing consumption patterns. The goal of the paper is to provide researchers with tools for the analysis of consumption. Our approach bridges the gap between detailed product-level analysis and broad consumption categories, providing a more granular understanding of consumer behavior. We apply the framework in several applications. Our empirical findings underscore the importance of product characteristics in shaping consumption trends. During the COVID-19 pandemic, the degree of human interaction required for consumption was a key determinant of expenditure patterns. Over the long run, within services, we find that the time investment required to use a product is an important predictor of expenditure growth.

Across the applications in the paper, the same message emerges: coarse expenditure categories miss economically important variation in what products are like and how they are consumed. By organizing goods and services in a common characteristic space, the ALTZ content model provides a way to describe, compare, and explain movements in demand that are difficult to see through standard classifications alone.

Our goal is for the content model and the underlying data to continue to grow and to be applied to a broader set of economic questions. For example, in the study of life-cycle expenditures, a natural application is to map expenditure profiles over the life cycle and across generations into the 23 characteristics. Another application is to use the framework to study fiscal transmission by asking which types of characteristics are most responsive to injections of liquidity to households.

References

Aguiar, M. and E. Hurst (2005). Consumption versus expenditure. *Journal of political Economy 113*(5), 919–948.

Aguiar, M. and E. Hurst (2007a). Life-cycle prices and production. *American Economic Review 97*(5), 1533–1559.

Aguiar, M. and E. Hurst (2007b). Measuring trends in leisure: The allocation of time over five decades. *The quarterly journal of economics 122*(3), 969–1006.

Bailey, K. D. (1994). *Typologies and taxonomies: An introduction to classification techniques*, Volume 12. Sage.

Baxter, M. and U. J. Jermann (1999). Household production and the excess sensitivity of consumption to current income. *American Economic Review 89*(4), 902–920.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal 75*(299), 493–517.

Benhabib, J., R. Rogerson, and R. Wright (1991). Homework in macroeconomics: Household production and aggregate fluctuations. *Journal of Political Economy 99*(6), 1166–1187.

Boppart, T. (2014). Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica 82*(6), 2167–2196.

Breiman, L. (2001). Random forests. *Machine learning 45*, 5–32.

Buera, F. J. and J. P. Kaboski (2012). The rise of the service economy. *American Economic Review 102*(6), 2540–2569.

Cirelli, F. and M. Gertler (2025). Economic winners versus losers and the unequal pandemic recession. *American Economic Journal: Macroeconomics 17*(3), 342–371.

Comin, D., D. Lashkari, and M. Mestieri (2021). Structural change with long-run income and price effects. *Econometrica 89*(1), 311–374.

Ford, G. T., D. B. Smith, and J. L. Swasy (1988). An empirical test of the search, experience and credence attributes framework. *Advances in Consumer Research 15*(1), 239–243.

Gilboa, I., A. Postlewaite, and L. Samuelson (2016). Memorable consumption. *Journal of Economic Theory 165*, 414–455.

Hai, R., D. Krueger, and A. Postlewaite (2020). On the welfare cost of consumption fluctuations in the presence of memorable goods. *Quantitative Economics 11*(4), 1177–1214.

Hendler, R. (1975). Lancaster's new approach to consumer demand and its limitations. *The American Economic Review 65*(1), 194–199.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2013). Two perspectives on preferences and structural transformation. *American Economic Review 103*(7), 2752–2789.

Lancaster, K. (1971). Consumer demand: A new approach.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of political economy 74*(2), 132–157.

Nickerson, R. C., U. Varshney, and J. Muntermann (2013). A method for taxonomy development and its application in information systems. *European journal of information systems 22*(3), 336–359.

Ramey, V. A. and N. Francis (2009). A century of work and leisure. *American Economic Journal: Macroeconomics 1*(2), 189–224.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy 82*(1), 34–55.

Taylor, L. D. and H. S. Houthakker (2009). *Consumer demand in the United States: Prices, income, and consumption behavior*. Springer Science & Business Media.

Appendix

A Additional Results

A.1 Explaining Price & Income Elasticities

Taylor and Houthakker (2009) is a standard reference for estimates of price and income elasticities for consumer products. They provide estimates for BEA-defined product categories disaggregated at the one-, two-, and three-digit levels (*e.g.*, Transportation, User-Operated Transportation, New Autos). In this section we use our framework to try to understand *why* expenditures on certain products are more or less responsive to changes in prices and income. Put differently, we estimate the extent to which high (or low) elasticity products from Taylor and Houthakker (2009) can be represented as a lower-dimensional combination of our CPCP set of product characteristics.

We use a stepwise regression model to select the characteristics most useful for explaining Taylor and Houthakker (2009)'s estimated price and income elasticities (taken from their Table 16.2). Our estimates are derived with a backward-selection criterion, starting with all characteristics and then progressively removing them if their p-value is above a threshold of 0.05.

Table 8: Stepwise regression estimates. Backward selection with threshold significance level of 0.05. Dependent variable: (1) Price elasticity, (2) Income elasticity.

	(1)	(2)
ConsFreq	0.304**	
	(0.007)	
Tangible	0.340**	
	(0.002)	
Habit	-0.263*	
	(0.019)	
Conspic		0.377***
		(0.001)
Subsist		-0.344**
		(0.001)

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Our results are reported in Table 8. They are intuitive. Products that tend to be more price elastic tend to be the ones that are purchased more frequently and are tangible. Products that induce a habit (*e.g.*, tobacco, alcohol) tend to be less price elastic. Income elasticities appear to be higher for conspicuous-consumption products, and lower for products that are subject to a lower bound on consumption required for the consumer's subsistence.

A.2 Robustness: Pandemic Drop and Recovery

Cobb-Douglas Benchmark The following tables replicate the results in Table 5 and Table 6 in the case of a Cobb-Douglas specification for preferences.

Table 9: Lasso estimates of β_j^c in equation (6) for the pandemic drop, April 2019 to April 2020, under the Cobb-Douglas benchmark ($\sigma = 1$). Columns report adaptive lasso and plugin lasso.

	Adaptive lasso	Plugin lasso
Conspic	-0.165	
Subsist	0.087	
Charity	0.148	
Tangible	-0.101	
Durable	0.235	
ConsFreq	0.150	
ConsHuman	-0.207	-0.386
Constant	-0.212	-0.212
Selected characteristics	7	1
λ	0.0749	0.4348

Table 10: Lasso estimates of β_j^c in equation (6) for the pandemic recovery, April 2020 to April 2021, under the Cobb-Douglas benchmark ($\sigma = 1$). Columns report adaptive lasso and plug-in lasso.

	Adaptive lasso	Plugin lasso
Conspic	0.092	0.183
Subsist	-0.099	
Charity	-0.058	
Tangible		0.132
Rival	0.150	
Durable	-0.115	
ConsHuman	0.154	
ConsInv	0.070	
Constant	0.164	0.164
Selected characteristics	7	2
λ	0.0199	0.4348

Correlation of Characteristics The correlation matrix across characteristics is displayed in Figure 9 below.

B Surveyed Products and Characteristics

In this section we list the set of products surveyed by the professional data firm *Civic Science* (CS) as described in Section 3.5.

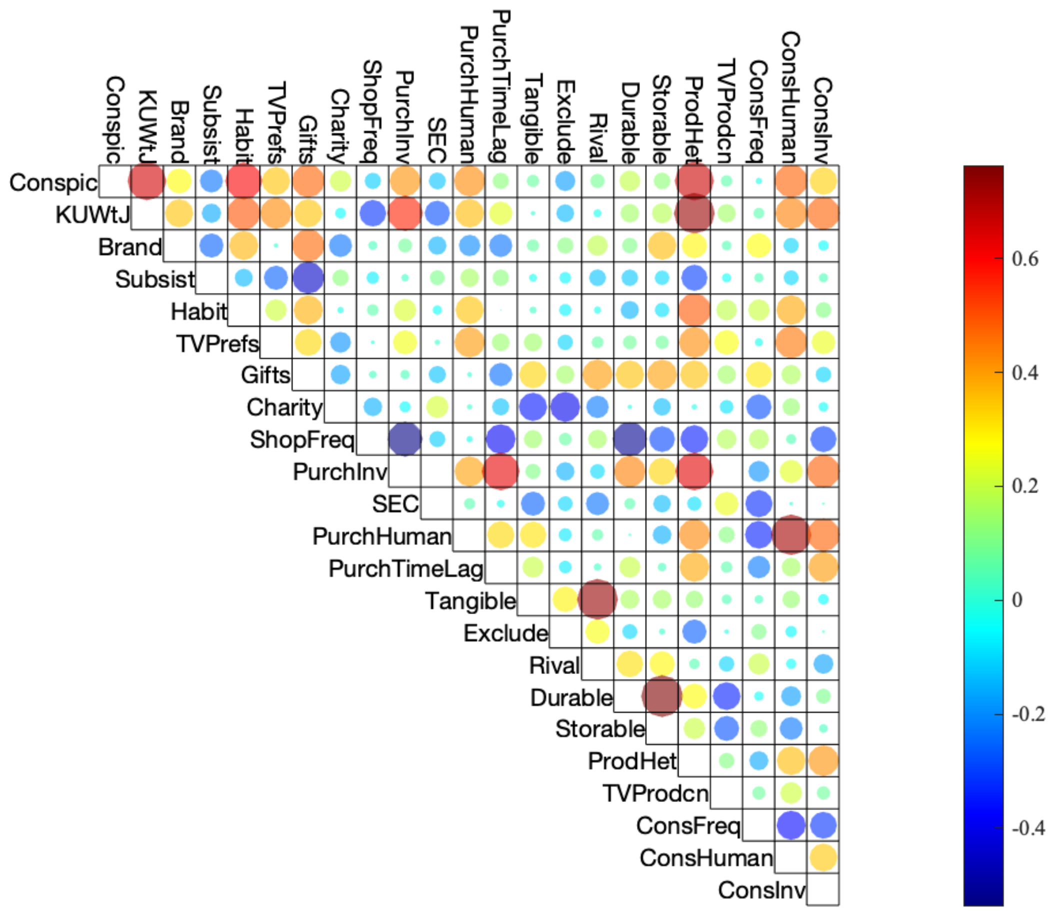


Figure 9: Cross Correlation of Characteristics in ALTZ content model.

Table 11: Surveyed Products.

Product/Service
Groceries
Clothes and shoes
Video or board games, toys, hobbies
New or used cars and trucks
Furniture
Audio and video equipment
Financial advisory or brokerage services
Health care (not in-hospital)
Hospital health care
Going to college or trade school
Legal or tax services
Auto repairs
Going to movies, shows, sporting events
Fast food
Going to a restaurant
Foreign travel (airfare, hotel, meals)
Housing (renting or owning)
Nursing-home care

C BEA Product List

The following tables 12, 13, and 14 list the 69 products considered in the ALTZ content model. The different products are grouped according to their respective BEA classification as durables, non-durables, or services.

Table 12: Product List: Part I (BEA Durable Goods).

ALTZ Product Number	Product Name
1	New and used autos and motorcycles
2	Motor vehicle parts and accessories
3	Furniture and household furnishings
4	Household appliances
5	Glassware, tableware, and household utensils
6	Tools and equipment for home and garden
7	Video and audio equipment
8	Computers/tablets/peripherals, computer software, other
9	Sporting equipment
10	Bicycles and bicycle accessories
11	Pleasure boats, aircraft, and other recreational vehicles
12	Recreational and Educational Books
13	Musical instruments
14	Jewelry and watches
15	Personal medical equipment, eyeglasses
16	Luggage
17	Telephones and communications

Table 13: Product List: Part II (BEA Nondurables Goods).

ALTZ Product Number	Product Name
18	Food and non-alc. beverages purch. from non-dining places
19	Alcoholic beverages for home, tobacco
20	Clothing and footwear
21	Gasoline, lubricants, fuel and other energy products
22	Pharmaceutical drugs (prescription and non-prescription)
23	Recreational items
24	Household supplies, disposable and somewhat durable
25	Personal care products
26	Magazines, newspapers, and stationery

Table 14: Product List: Part III (BEA Services).

ALTZ Product Number	Product Name
27	Housing Services
28	Household utilities
29	Outpatient Health Care Services
30	Hospital Health Care Services
31	Nursing homes
32	Motor vehicle maintenance and repair
33	Motor vehicle leasing and rental
34	Parking fees and driving tolls
35	Public land transportation
36	Air transportation
37	Water transportation
38	Recreation services
39	Television services
40	Photo and audiovisual services
41	Video and audio streaming and rental
42	Gambling
43	Pet services
44	Package tours
45	School lunches and food supplied by government
46	Meals at limited service eating places
47	Meals at full-service restaurants, trains, movies, perf. arts, arenas
48	Alcohol at Eating/Drinking Places
49	Accommodations in hotels and motels
50	Accommodations at schools
51	Retail banking services
52	Financial services other than retail banking
53	Insurance
54	Telecommunication services
55	Postal and delivery services
56	Internet access
57	Educational services, higher education and vocational
58	Educational services, elementary and secondary schools
59	Legal services, accounting services, other business services
60	Labor and professional organization dues
61	Funeral and burial services
62	Personal care services
63	Clothing and footwear services
64	Child care
65	Day care and nursery schools
66	Social assistance
67	Giving to social advocacy and religious institutions
68	Household maintenance
69	Foreign travel by US residents