Submitted to *Management Science*
Initial Submission: March, 2025

Decoding Consumer Preferences: Using Attention-Based Language Models for Demand Estimation and Counterfactual Analysis

(Authors' names blinded for peer review)

This paper proposes a new demand estimation method using attention-based language models. We train an encoder language model in a two-stage process to analyze the natural language descriptions of used cars from a large US-based online auction marketplace. Our approach enables us to semi-nonparametrically estimate the demand primitives of our structural model (private valuations and market size) for each vehicle in the dataset. In the first stage, we fine-tune the language model to encode the target auction outcomes using the natural language descriptions provided for each vehicle. In the second stage, we project the trained language model's encodings into the parameter space of two density estimators: market size and consumer valuation. The model's capability to conduct counterfactual analyses within the trained market space is validated using a subsample of withheld auction data, which includes a set of unique "zero shot" instances.

Key words: Consumer Preferences; Demand Estimation; Large Language Models; Semi-Nonparametric Estimation; Hedonic Value

1. Introduction

Identifying the determinants of consumer value is a fundamental challenge of sales. For many products, commercial success hinges on how well their features reflect the underlying consumer preferences in the market, and thus firms have a strong economic incentive to determine how various combinations of features produce value. One common approach to estimating the strength of consumer preferences is to convert raw, unstructured feature data into a structured dataset that could plausibly delineate the sources of all value. For instance, in a used car market one might collect data on vehicles' make, model, year, and miles, while for theater shows one might collect data on the type of show, cast, venue seat, and day-of-week. Using these data, prices can be modeled as a function of the overall value, as estimated by summing the *hedonic* values of the features in the product (Rosen 1974).

As a generic illustration, consider a standard approach presented in Figure 1. First, a set of features X is defined from a raw, unstructured data source (e.g., vehicle description in a used car market) to serve as inputs into a structural estimation model $\mathcal{F}$. The structural model produces estimates $\hat{Y}$ on

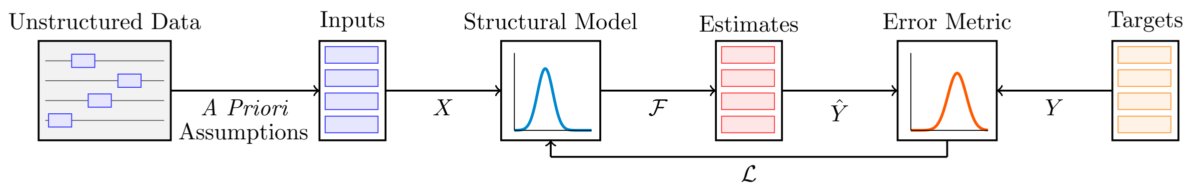


Figure 1 Standard Estimation Model

targets Y (e.g., the sales price of used cars), and an error metric $\mathcal{L}$ is used to optimize the parameters of the structural model.

Although simple and intuitive, a drawback to this approach is that it often involves making *a priori* assumptions on the sources and structure of consumer valuations. Without sufficient knowledge of how to code the true sources of value, one must simply rely (perhaps unscientifically) on intuition for this task (Ludwig and Mullainathan 2024). And, even with perfect knowledge of the feature set, feature-based preferences can be diverse, dynamic, and interdependent – potentially requiring complex identification methods in order to estimate the true hedonic values of consumers (Cropper, Deck and McConnell 1988). Furthermore, any resulting insights will be limited to the defined attributes and specific market dynamics, as it is not immediate how, for instance, a standard regression-based inference might apply to an out-of-sample good, or a set of previously unobserved attributes, or under different market mechanisms.

Motivated by the above limitations and the recent development and applications of natural language processing, we offer a new approach to modeling the sources of consumer valuations. Specifically, we feed a detailed text description of a good within a specific market environment to an attention-based language model, which converts the text into *embedding* vectors of a high-dimensional, real-valued space that can numerically represent the good's underlying demand and market information. We label this space as the *Value Embedding Space*, and each specific embedding vector, corresponding to a partial piece of the description of a particular good, as *Demand Embedding Vectors* (DEVs). The benefit of this language-model-based approach is that it can process and interpret intricate language patterns, contextual information, and nuanced connotations of product features, and thus generate more complex representations of consumer preferences.

The proposed approach employs a two-stage estimation process, as illustrated in Figure 2. In Stage 1, a language model $\mathcal{M}$ is trained to generate DEVs that *encode* point predictions $\hat{Z}$ for the target outcomes Z using textual descriptions of the products of a market as input. To achieve this, a weight vector $\mathcal{H}_1$ is simultaneously optimized to project the DEVs into the prediction space corresponding to these outcomes. For the empirical illustration of online car auction data, the target outcomes include metrics on bid values, the number of unique bidders, auction views, watchers, and whether

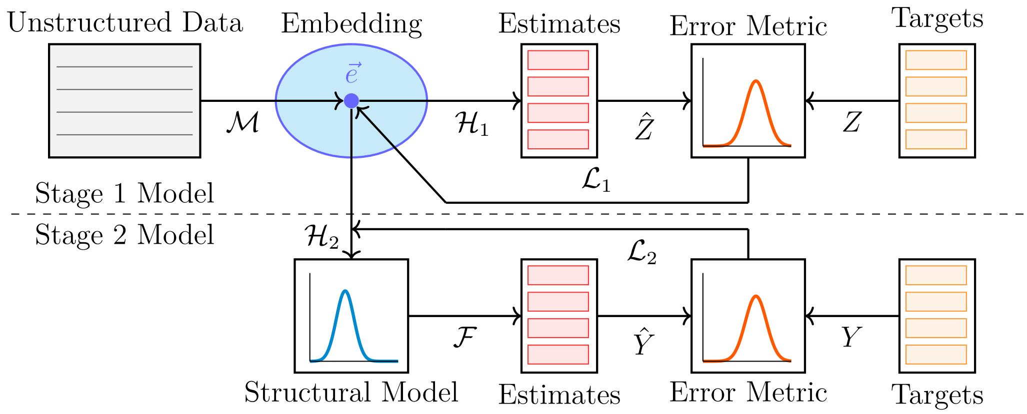


Figure 2 Proposed Model

the reserve price was met. Both $\mathcal{M}$ and $\mathcal{H}_1$ are trained with an error metric $\mathcal{L}_1$, ensuring that the DEVs effectively capture the necessary market information to serve as the foundation for subsequent econometric analysis.

In Stage 2, a fully connected feedforward network, $\mathcal{H}_2$, *decodes* the market information embedded in the DEVs from Stage 1 into the parameter space of the structural model $\mathcal{F}$, which produces estimates $\hat{Y}$ for the target market outcomes Y, which may be a subset of the target market outcomes Z from Stage 1. This second stage is crucial for enabling robust counterfactual analysis, as it explicitly constrains the model to use the assumed data generating process to predict the observed target outcomes. In the empirical illustration, these estimates are obtained using semi-nonparametric methods to model the probability density functions of consumer valuations and market size for each auction vehicle's description. The network $\mathcal{H}_2$ is trained with an error metric $\mathcal{L}_2$, optimizing over the predicted outcomes from $\mathcal{F}$ while ensuring adherence to the specified structural relationships.

This *encoder-decoder* approach compresses relevant market information into DEVs during Stage 1 and then transforms it into optimized estimates within the econometric framework in Stage 2. By fine-tuning the language model in Stage 1, the model learns subtle patterns in product descriptions that reflect consumer preferences and the value generated by these products, establishing a sophisticated mapping between unstructured textual features and market outcomes. The structural constraints imposed in Stage 2 then ensure that counterfactual analyses remain grounded in economic theory, making the predictions more reliable when simulating market interventions or policy changes. This approach overcomes key limitations of traditional hedonic models by eliminating the need for *a priori* assumptions, enhancing flexibility and scalability across diverse markets and product categories, and improving the granularity of analysis to detect nuances in consumer preferences and market dynamics.

To illustrate the proposed model, we collected transactional data from a large US online auction house for used vehicles. Using time-stamped and user-identifiable bids on approximately 80,000 car auctions from 2014-2023, we estimate two primitives of demand for our structural model: the distribution of bidders' private valuations and the distribution of the number of potential bidders (i.e., market size). To enrich the diversity of training data provided to our model, we also tasked a generative AI model with producing nearly 300,000 synthetic descriptions of the vehicles using expert personas that could evaluate the vehicles from their description (an independent vehicle history report provider; a certified automotive mechanic; a car enthusiast club president; and a vehicle appraiser). To validate the final model's ability to generalize to previously unseen descriptions, we used a holdout sample of 5,000 authentic auction observations from our dataset, including an ablation study, or "zero shot" test, for a specific make and model of vehicle – the DMC DeLorean.

Our ability to offer these results is based on two recent developments. First, the scaling of machine learning algorithms has generated powerful emergent abilities to parse semantic meaning from natural language. In the context of natural language processing, an attention-based language model's architecture allows it to establish nonlinear relationships between various components of a text sequence (Mikolov, Yih and Zweig 2013, see Qiu et al. (2020) for a survey). After being trained on large swaths of the Internet to predict missing or "masked" text elements, these models can later be fine-tuned for alternative natural language applications (Devlin et al. 2018). If tasked with predicting market outcomes such as prices or bids from the natural language descriptions of products, as we do here, training the language model will convert its initial semantic representations of goods to demand representations that are suitable for downstream econometric estimation. Second, we benefit from an abundance of naturally occurring market data, such as those in our online auctions, that can offer strong signals of individual consumer valuations while also satisfying the training requirements of a language model.

The managerial implications of integrating machine learning with economic modeling extend beyond improved prediction – they redefine how firms generate causal insights for strategic decision making (Athey 2018). Language models unlock vast new data sources, such as product descriptions and written consumer feedback. However, their outputs must be structured within econometric models that enable reliable counterfactual analysis. The contribution of this paper is that our approach bridges the aforementioned gap by embedding language-model-driven demand signals into a structural model, and thus allowing managers to move beyond correlation-based forecasting to derive actionable, policy-relevant insights. This is particularly critical in digital marketplaces, where firms must anticipate demand shifts, optimize platform design, and refine dynamic pricing strategies in response to evolving market conditions. By integrating machine learning representations with economic inference techniques, our model ensures that machine-driven recommendations align with underlying economic forces, enhancing their applicability for real-world inferences. However, realizing this potential requires overcoming key challenges, including bias mitigation, interpretability, and model robustness across different market contexts (Van Giffen, Herhausen and Fahse 2022). As firms increasingly rely on machine learning for decision support, our approach offers a framework for merging algorithmic flexibility with economic inference, establishing a foundation for machine-powered business strategy.

The rest of the paper proceeds as follows. Following an overview of the existing literature in Section 2, we provide a more detailed description of the model framework in Section 3, and the technical model training in Section 4. In Section 5 we provide an overview and summary statistics of the online auction data, and then in Section 6 provide an empirical illustration of the proposed model framework, and discuss the numerical results and counter-factual analysis from the online auction data. Finally, in Section 7, we conclude with some remarks and ideas for future extensions.

Throughout this paper, we adhere to a notational convention in which uppercase letters indicate random variables, lowercase letters represent their realizations, Greek letters are used to denote model parameters, and cursive letters denote a model-specific function. As the ensuing empirical illustration is based on vehicles, we use the term "product" instead of "good," but stress that the presented framework equally applies to "services".

2. Literature Review

With the creation of the transformer architecture (Vaswani et al. 2017), attention-based language models have achieved unprecedented capabilities in representing the semantic meaning of natural language. Initially developed for natural language translations, these models have since demonstrated broad applicability across diverse domains, including text generation (Radford et al. 2019), protein structure prediction (Jumper et al. 2021), medical image processing (Chen et al. 2021), writing advertising copy (Chen and Chan 2024), privacy policy changes (Lin et al. 2024), and prediction on course and instructor evaluation (Wang et al. 2025). More recently, these advancements have led to a surge in the use of language models for empirical research in economics, business, and management, where textual data is increasingly leveraged to make economic insights on consumer preferences, firm strategies, and market and business dynamics (Gentzkow, Kelly and Taddy 2019, Berger et al. 2020, Ash and Hansen 2023, de Kok 2025).

Although language models have proven effective in areas such as sentiment prediction in customer service (Puranam, Kadiyali and Narayan 2021), product reviews (Ma and Luo 2023), and financial disclosures (McCarthy and Alaghband 2024), their potential for demand estimation remains underexplored. Recent advances suggest that language models may provide a useful tool for extracting latent demand signals from natural language descriptions of goods and services, offering new possibilities for new approaches to demand estimation in a variety of market settings (Timoshenko and Hauser 2019, Wang et al. 2023, Liu 2023, Wang, Zhang and Zhang 2024).

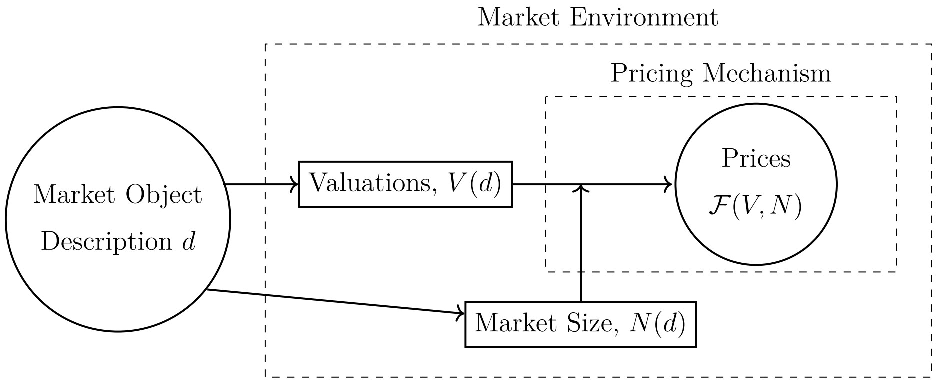


Figure 3 **Assumed Data Generating Process**
Note: Market observables are in circles, latent variables are in rectangles.

Furthermore, the integration of machine learning into econometric methods has led to advances in predictive accuracy and modeling flexibility, particularly in settings where traditional methods face limitations (Athey and Imbens 2019). Recent research has leveraged deep learning techniques to capture complex relationships between product attributes and consumer preferences, incorporating high-dimensional representations derived from structured and unstructured data sources (Adam, He and Zheng 2024, Aceves and Evans 2024). Language models, in particular, have been used to explore individual preferences and risk tolerances in a variety of settings (Dillion et al. 2023, Wu et al. 2023, Goli and Singh 2024, Zhu, Yan and Griffiths 2024, Netzer, Lemaire and Herzenstein 2019).

Although these methods improve predictive performance, a key challenge remains their lack of structural coherence and interpretability.[1] Models that directly estimate economic primitives within a structural framework are typically better suited for counterfactual analysis (Verma et al. 2024), and improve performance on "zero-shot" (i.e. novel) problems (Wei et al. 2021), yet many machine learning approaches lack this foundation. Recognizing this, recent work has explored hybrid approaches that integrate machine learning representations into structural models, using deep embeddings as inputs to traditional econometric frameworks (e.g. Vafa, Athey and Blei 2024). These approaches suggest a path forward for demand estimation methods that combine the flexibility of machine-driven representations with the interpretability and counterfactual validity of structural econometric models.

3. Econometric and Machine Learning Framework

The framework for our proposed model begins with a set of assumptions over the data generating process on market demand, outlined in Figure 3. Within a given market environment, the introduction of a product with description d generates, to an exposed population, two stochastic demand primitives that together produce an aggregate demand schedule. The first primitive of interest is

[1] See Olah et al. (2020), Elhage et al. (2022), Templeton (2024) for language model interpretability methods.

Table 1 Pedagogical Example of Tokenization and Demand Embedding Vectors

Description d	This	1974	BMW	3	.	0	CS	was	refurb	ished	...
Tokens $\mathcal{D}: d \to \vec{\mathsf{d}}$	713	15524	8588	155	4	288	6842	21	17880	6555	...
DEVs $\mathcal{M}: \vec{\mathsf{d}} \to \{\vec{e}_j\}$	$\vec{e}_1$	$\vec{e}_2$	$\vec{e}_3$	$\vec{e}_4$	$\vec{e}_5$	$\vec{e}_6$	$\vec{e}_7$	$\vec{e}_8$	$\vec{e}_9$	$\vec{e}_{10}$	...

$V(d)$, a continuous random variable that produces *i.i.d.* realizations on the *willingness to pay* that consumers have for d. The corresponding distribution function is given by $F_V(v) = \Pr\{V(d) \leq v\}$, with support on $\mathbb{R}$. The second primitive of interest is $N(d)$, a discrete random variable that represents the *number of potential consumers* a description d attracts (i.e. market size). Likewise, this is defined by the distribution function $F_N(n) = \Pr\{N(d) \leq n\}$, with support on $\mathbb{N}$. To generate a demand schedule, we assume that the realization in market size n determines the set of realizations over valuations $\{v_1, v_2, \ldots, v_n\}$. Without loss of generality, we assume that the valuations are ordered, i.e. $v_n \leq \cdots \leq v_2 \leq v_1$. For simplicity, we assume that all consumers have unitary demand, and that $V(d)$ and $N(d)$ are independent for all d, i.e. $f_{V,N}(v,n) = f_V(v) f_N(n)$.

Once valuations are realized, consumers engage with the market's pricing mechanism, which can be a function of the market structure and available information (as an example) to produce observable prices. In the context of our empirical application of online vehicle auctions, our structural model $\mathcal{F}$ follows the format of an English auction, which implies that the demand for a vehicle is stochastically determined when it is up for auction. We do not predetermine the dynamics of $V(d)$ and $N(d)$ or the conditions for their realization. Instead, we leverage the language model to uncover these dynamics based on the vehicle's description, imposed structural model, and available market data. Next, we discuss the two-stage estimation method for the distributions of V and N, as presented in Figure 2.

3.1. Encoding Economic Information in Market Object Descriptions

The proposed estimation approach uses a language model to estimate the demand characteristics of a market object from the features represented in its description d. For this task, d is broken down into a vector of the language model's most basic unit of text, called a *token*, which we denote as d. Tokens are numerical representations of words, subwords (words broken into smaller parts), or characters. To convert text to tokens, the language model is accompanied by a designated *tokenizer*, which is a function $\mathcal{D}$ that converts a text string to a vector of integer values within the language model's vocabulary set, $\mathcal{D}: d \to \vec{\mathsf{d}}$. We denote the length of this token vector as $\mathsf{d}_{\max}$. A pedagogical example of tokenization is provided in Table 1 using the tokenizer we employ in the empirical section of this study, which has a vocabulary set of approximately 50,000 tokens (Liu et al. 2019).

Attention-based language models are trained to process tokens in a way that encodes the information contained within the represented text. In our application, the specific target data that we train the language model to encode are the relevant and observable market demand information available on the object. This implies that each token d_j is mapped to a q-dimensional real-valued encoding vector $\vec{e_j} \in \mathbb{R}^q$, $j = 1, \ldots, \mathsf{d}_{\max}$, labeled *Demand Embedding Vector*. In other words, for a given description d, converted into a vector of tokens $\vec{\mathsf{d}}$, the language model $\mathcal{M}$ maps the sequence of tokens to a corresponding sequence of DEVs (see Table 1),

$$\mathcal{M} : \vec{\mathsf{d}} \rightarrow \{\vec{e_j}\}_{j=1}^{\mathsf{d}_{\max}}$$

One of the key features of how modern language models process tokens to produce embeddings is their self-attention mechanism. This mechanism allows the model to determine the relevance of each token in a text based on its relationship with all other tokens (see Turian, Ratinov and Bengio 2010, Mikolov, Yih and Zweig 2013, Vylomova et al. 2015). This means that language models do not assign a fixed encoding to each token within its vocabulary set. Rather, the model calculates a weighted sum of all tokens in the sequence, where the weights reflect the relevance of each token to the one being encoded. As an example, the DEVs assigned to the tokens for "manual transmission" will be dependent on the other tokens in $\vec{\mathsf{d}}$, such as whether those representing "Honda" or "Lamborghini" are present. When paired with fine-tuning methods, this process enables attention-based language models to generate detailed and context-aware representations of tokens, which is critical to producing representative DEVs (Li et al. 2020).

An important yet perhaps non-obvious design decision for our application of a language model is the total number of tokens processed with each description. If object descriptions produce token sequences of various lengths, then the language model will not be capable of generating consistent estimates of a token's contribution to the target variable, even when two tokens have identical demand vectors. Consequently, in the ensuing empirical analysis, while training the model we fix the number of tokens the language model processes in every auction to have the same value $\mathsf{d}_{\max}$. In instances where the tokenizer receives a description that produces more or fewer tokens than $\mathsf{d}_{\max}$, the description is truncated or padded, respectively, to maintain the fixed sequence length. The empirical analysis on online vehicle auctions is based on a token embedding space of $q = 768$ and a token sequence length of $\mathsf{d}_{\max} = 512$.

3.2. Stage 1 Estimation: Token Level Contributions to Market Outcomes

The goal of the first stage estimation is to (a) train $\mathcal{M}$ to *encode* the representation each token $\mathsf{d}_j \in \vec{\mathsf{d}}$ has as a Demand Embedding Vector $\vec{e_j}$, given the product description d, and, simultaneously, (b) train a fully connected network head $\mathcal{H}_1$, appended to $\mathcal{M}$, to project each DEV $\vec{e_j}$ to m real-valued target outcomes of interest, $\hat{y}_j \in \mathbb{R}^m$. For example, in the case of a vehicle auction, these outcomes could include the submitted bid values, the number of active bidders, and how many times the auction was viewed. This approach leads to the following proposition, which reformulates a well-established result from the computer science literature in the context of our economic application.

Table 2 Schematic Illustration of Stage 1 Model Making m Point-predictions

Targets	$\{y_1(d),$	$\cdots$	$y_i(d),$	$\cdots$	$y_m(d)\}$
Token 1	$\{\hat{y}_{1,1}$	$\cdots$	$\hat{y}_{i,1}$	$\cdots$	$\hat{y}_{m,1}\}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
Token j	$\{\hat{y}_{1,j},$	$\cdots$	$\hat{y}_{i,j},$	$\cdots$	$\hat{y}_{m,j}\}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
Token $\mathsf{d}_{\max}$	$\{\hat{y}_{1,\mathsf{d}_{\max}},$	$\cdots$	$\hat{y}_{i,\mathsf{d}_{\max}},$	$\cdots$	$\hat{y}_{m,\mathsf{d}_{\max}}\}$
Predictions	$\{\hat{y}_1(d)=\sum_{j=1}^{\mathsf{d}_{\max}}\hat{y}_{1,j}$	$\cdots$	$\hat{y}_i(d)=\sum_{j=1}^{\mathsf{d}_{\max}}\hat{y}_{i,j}$	$\cdots$	$\hat{y}_m(d)=\sum_{j=1}^{\mathsf{d}_{\max}}\hat{y}_{m,j}\}$

Proposition 1. Let $D=\{(d,\vec{y})\}$ be a set of text descriptions d and associated economic information vectors $\vec{y}\in\mathbb{R}^m$. For any error tolerance $\epsilon>0$, there exists a language model $\mathcal{M}: D\to\mathbb{R}^{q\times d_{max}}$, and a decoding function $\mathcal{H}_1:\mathbb{R}^{q\times \mathsf{d}_{\max}}\to\mathbb{R}^m$, such that

$$\sup_{(d,y)\in D}\|y-\mathcal{H}_1\circ\mathcal{M}\|_2<\epsilon.$$

Proof. *See Appendix A*

Drawing on established universal approximation theorems in neural networks, this result simply restates that any information purposefully embedded in the language model's representation space is recoverable with arbitrary precision through an appropriate decoding function. Therefore, we apply this approach to generating predictions $\hat{y}_{i,j}$ on each token j's contribution to target outcome i, and produces a set of m observed market outcomes $\{y_1(d),y_2(d),\ldots,y_m(d)\}$ with the estimated decoder,

$$\hat{\mathcal{H}}_1:\vec{e}_j\to\{\hat{y}_{i,j}\}\quad\text{for all } i\in\{1,\ldots,m\}, j\in\{1,\ldots,\mathsf{d}_{\max}\}. \tag{1}$$

That is, the purpose of the $\hat{\mathcal{H}}_1$ head is to *decode* the high-dimensional DEV representations of the tokens in the object's description into their individual contributions to the target outcomes. This approach takes inspiration from the SHAP Value framework for additively estimating feature importance (Lundberg and Lee 2017), therefore we produce aggregate predictions for each $\hat{y}_i$ by summing the individual token predictions on market outcome i together,

$$\hat{y}_i=\sum_{j=1}^{\mathsf{d}_{\max}}\hat{y}_{i,j}\quad\text{for all } i\in\{1,\ldots,m\}$$

A schematic illustration of how token-level contributions produce estimates on the target outcomes is provided in Table 2.

We optimize the estimated parameters in $\hat{\mathcal{M}}$ and $\hat{\mathcal{H}}_1$ with a loss function that minimizes the mean squared error (MSE) between the estimated market outcomes $\{\hat{y}_1(d),\hat{y}_2(d),\ldots,\hat{y}_m(d)\}$ and the true observed market outcomes $\{y_1(d),y_2(d),\ldots,y_m(d)\}$. This is defined as

$$\mathcal{L}_1=\frac{1}{|D|}\sum_{d\in D}\sum_{i=1}^{m}\left[y_i(d)-\hat{y}_i(d)\right]^2 \tag{2}$$

where D represents the set of training samples. The loss function $\mathcal{L}_1$ ensures that the model learns to produce accurate point predictions by reducing the squared deviations between the predicted and actual market outcomes. We label the ensuing models, with the optimized parameters minimizing $\mathcal{L}_1$, by $\hat{\mathcal{M}}^*$ and $\hat{\mathcal{H}}_1^*$.

While Stage 1 provides valuable insights into token-level contributions, it does not yet offer direct inference on the underlying structural parameters governing market outcomes. To address this, the second stage estimation process leverages the DEVs to project market-relevant information into the parameter space of an econometric model, allowing for a more interpretable and structured analysis of demand primitives. We describe this process in the next subsection.

3.3. Stage 2 Estimation: Token Level Contributions to Demand Primitives

The objective of Stage 2 is to estimate two latent demand primitives of our structural model: the valuation distribution $F_V(v)$ and the market size distribution $F_N(n)$. While the machine learning approach in Stage 1 successfully encodes market outcomes into DEVs derived from product descriptions, these predictions alone do not identify the structural economic forces driving demand; for further discussion, see Mullainathan and Spiess (2017), Iskhakov, Rust and Schjerning (2020). By estimating the respective demand primitives via a structural model, we constrain the prediction process to follow the assumed data generating process, which is essential for performing robust counterfactual analyses. This structural approach enables us not only to infer the underlying relationships generating consumer behavior but also to make reliable predictions for unseen applications, such as new product descriptions or market structures, by ensuring that all counterfactual scenarios remain consistent with the economic theory embedded in our model.

To enable a general and flexible representation of the valuation distribution, we employ a seminonparametric approach developed by Gallant and Nychka (1987), Fenton and Gallant (1996*a*), and Fenton and Gallant (1996*b*). Specifically, the density function $f_V(v)$ is approximated by,

$$\hat{f}_V(v) = c\left[1 + \sum_{k=1}^{\kappa} \alpha_k \left(\frac{v-\mu}{\sigma}\right)^k\right]^2 \varphi(v \mid \mu, \sigma) \tag{3}$$

where $\varphi(v \mid \mu, \sigma)$ is the Gaussian density component that imparts location (μ) and scale (σ) properties, and, for a given level $\kappa \in \mathbb{N}$, a Hermite polynomial expansion capturing the deviations from normality through the weighted terms α_k. The squaring of the expansion guarantees non-negativity, and the normalizing constant,

$$c = \left[\int_{-\infty}^{\infty} \left(1 + \sum_{k=1}^{\kappa} \alpha_k \left(\frac{z-\mu}{\sigma}\right)^k\right)^2 \varphi(z \mid \mu, \sigma) dz\right]^{-1}$$

ensures that $\hat{f}_V(v)$ integrates to one, and thus a well-defined probability density function. This specification yields a parameter space for the valuation distribution of $\vec{\lambda}_V = \{\alpha_1, \ldots, \alpha_\kappa, \mu, \sigma\}$.

For the market size distribution, we adopt a semi-nonparametric approach that uses the softmax function to construct a probability mass function. Specifically, the parameter space for the market size distribution $\vec{\lambda}_N = \{\rho_{\underline{n}}, \ldots, \rho_{\overline{n}}\}$, consists of likelihood values ρ_r on the discretized grid of potential bidder counts $B = \{\underline{n}, \ldots, \overline{n}\}$, $B \subseteq \mathbb{N}$. The likelihood values are then transformed via the softmax function, which exponentiates and normalizes them so that they are nonnegative and sum to one. This transformation yields a probability mass function for each n in the defined grid. That is, the probability mass function $f_N(n)$ is approximated by, for $n \in \{\underline{n}, \ldots, \overline{n}\}$,

$$\hat{f}_N(n) = \frac{\exp(\rho_n)}{\sum_{r \in B} \exp(\rho_r)} \tag{4}$$

This formulation not only captures the inherently discrete nature of market participation but also allows for a flexible, data-driven estimation of the underlying bidder distribution.

Following Farrell, Liang and Misra (2020) on deep network architecture, our Stage 2 estimation approach integrates a neural network with semi-nonparametric density estimation techniques. Specifically, we append a fully connected network head, $\mathcal{H}_2$, to the fine-tuned language model $\mathcal{M}^*$ from Stage 1. For each product description d, $\mathcal{M}^*$ encodes the set of token DEVs, $\{\vec{e}_j\}_{j=1}^{\mathsf{d}_{\max}}$, which are subsequently processed by $\mathcal{H}_2$. This network head maps the DEVs into the parameter spaces of both distributions,

$$\mathcal{H}_2 : \{\vec{e}_j\}_{j=1}^{\mathsf{d}_{\max}} \to [\vec{\lambda}_V, \vec{\lambda}_N].$$

The dual-head architecture ensures that although both distributions are derived from the same set of DEVs, each branch is tailored to its respective estimation task. The valuation branch outputs the Hermite polynomial coefficients along with Gaussian parameters to capture continuous consumer valuations, while the market size branch produces discrete probabilities for bidder participation via the softmax function. By mapping DEVs into these conditional density spaces, $\mathcal{H}_2$ effectively translates the market-relevant signals encoded in Stage 1 into coherent, interpretable demand estimates that are well-suited for structural analysis.

In order to train $\mathcal{H}_2$, i.e. estimate the model parameters, we need to specify how the latent demand primitives, through the specific pricing mechanism P from Figure 3, generate the observed target outcomes. In other words, the calibration of $\mathcal{H}_2$ is context-dependent on the data generating process. Motivated by the ensuing empirical analyses, we illustrate by considering the English auction format as the pricing mechanism.

The English Auction Pricing Mechanism and Equilibrium Bidding

In an English auction participating bidders sequentially submit incremental bids until no bidder is willing to exceed the current highest bid and the auction concludes. It should be fairly intuitive that, a rational bidder l, with private valuation v_l, should continue to incrementally bid b_l^+, as long as their surplus $v_l - b_l^+$ is non-negative. Hence, bidders' weakly dominant strategy is to bid up to their valuation, $b_l^+ \le v_l$, and, conditional on the seller's reserve price p_R being met, the final transaction price is determined by the *second highest valuation* plus the bid increment submitted by the winning bidder. To simplify notation, denote bidder l's *last bid* as b_l, $l = 1, \dots, n$. Given ordered valuations, it follows that $b_l = v_l$, for $l = 2, \dots, n$, while $b_1 = v_2 + b_{inc}$, where b_{inc} is the minimum bid increment. That is, for a subset of bidders, namely all non-winning bidders, we assume the last (maximum) bid they submit represents their valuation. Note, in the English auction the winning bidder's valuation is censored and only known to exceed the second highest valuation with the minimum bid increment. We apply this revelation of valuations through the equilibrium bid strategy to the auction data.

While the equilibrium bidding strategy for the English auction is robust to several complicating factors, including risk tolerance and the number of bidders (Wolfstetter 1996), there are a couple of potential shortcomings. First, it is unable to point identify the necessary order statistics in the presence of jump bidding, i.e. bidding much more than the minimum bid increment; see Avery (1998), Easley and Tenorio (2004). If there are large bid gaps due to jump bidding then assuming $b_l^* = v_l$ will produce biased estimates. Second, this is also a concern if participation is temporally intermittent, or sparse, or subject to rapid counter-bidding. For instance, if two bidders aggressively counter-bid each other, then the price trajectory might escalate quickly such that other bidders' last bids are far below their respective valuations. Finally, if there are multiple ongoing, or sequential, auctions of identical or similar objects, then bidders may strategically not bid up to their valuation in a given auction, in anticipation of maximizing their expected surplus over a set of auctions. These shortcomings may thus introduce estimation bias on the order statistics of the $2^{nd}, 3^{rd}, 4^{th}, \dots$ highest valuations. We address these concerns in our inclusion criteria for the auction data discussed in the empirical section of the paper.

Given that the equilibrium bid strategy can serve as reasonable proxies for bidders' private valuation, we use the approximated density and probability mass functions, $\hat{f}_V(v)$ and $\hat{f}_N(n)$, to predict the order statistics of the observed bids in the auctions. Specifically, recall that the number of bidders n is drawn from $F_N(n)$, and bidders' valuation are drawn i.i.d. from $F_V(v)$. This implies that conditional on a realized market size n, the estimated density function for the order statistics of the top $l \le n$ valuations, is given by, for $l' = l, \dots, 2, 1$,

$$\hat{f}_{V_{(l')}}(v|n) = \frac{n!}{(l'-1)!(n-l')!} \hat{f}_V(v)(\hat{F}_V(v))^{n-l'}(1-\hat{F}_V(v))^{l'-1} \tag{5}$$

Using Equation (5), the structural model $\mathcal{F}$ numerically computes the (conditional on market size n) expectation of the top l order statistic by,

$$\hat{E}\left[V_{(l')} \mid n\right] = \int_{-\infty}^{\infty} v \hat{f}_{V_{(l')}}(v|n)\, dv$$

To compute the unconditional expectation, we average over the distribution of n, for $l' = l, \ldots, 2, 1$,

$$\hat{E}_d\left[V_{(l')}\right] = \sum_{n=\underline{n}}^{\overline{n}} \hat{f}_N(n)\, \hat{E}\left[V_{(l')} \mid n\right]$$

Finally, we collect these expectations into a vector representing the predictions of the l largest bids,

$$\left[\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_l\right] = \left[\hat{E}_d\left[V_{(1)}\right], \hat{E}_d\left[V_{(2)}\right], \ldots, \hat{E}_d\left[V_{(l)}\right]\right] \tag{6}$$

To train $\mathcal{H}_2$, we optimize its parameters to ensure that the estimated valuation and market size distributions generate expected bid predictions via $\mathcal{F}$ that align with the observed bid results. As previously discussed, the largest bid b_1 is biased downward due to the bidding strategies induced by the English auction. Therefore, $\mathcal{H}_2$ is trained to estimate valuations $\{V_2, \ldots, V_l\}$, using the bids $\{b_2, \ldots, b_l\}$, according to their computed order statistics from $\mathcal{F}$. We do this by minimizing the mean squared error (MSE) between the predicted bids $\hat{b}$ and the observed bids b for the bids that can proxy for bidder valuations. The loss function is given by,

$$\mathcal{L}_2 = \frac{1}{|D|} \sum_{d \in D} \sum_{l'=2}^{l} \left[b_{l'}(d) - \hat{b}_{l'}(d)\right]^2 \tag{7}$$

where D represents the set of training auctions. This objective ensures that the estimated distributions $\hat{F}_V(v)$ and $\hat{F}_N(n)$ generate bid predictions that match empirical observations. Similar to above in Stage 1, we denote $\mathcal{H}_2^*$ as the model with the optimized parameters that minimize $\mathcal{L}_2$.

Although the estimation framework for $\mathcal{F}$ is tailored for English auctions, our broader approach can accommodate alternative market settings. For example, in posted-price markets, firms set prices based on expected residual demand rather than competitive bidding. Similarly, in bargaining environments, price formation is influenced by strategic negotiation between buyers and sellers. To extend our methodology to such settings, the structure of $\mathcal{F}$ must be adjusted to reflect the equilibrium conditions that govern the formation of prices in these alternative markets.

4. Language Model Training and Structural Model Estimation

The foundation of our training framework is the Robustly Optimized BERT Pretraining Approach (RoBERTa) language model introduced by Liu et al. (2019). RoBERTa is an advanced variant of the Bidirectional Encoder Representations from Transformers (BERT) model that demonstrated how a transformer-based language model could encode information from unstructured text (Devlin et al. 2018). In our application to online auction data, we use RoBERTa to initialize our language model $\mathcal{M}$, which facilitates a two-stage training process to extract structured demand representations and estimate the underlying economic primitives governing the observed auction outcomes. Table 3 summarizes training setup and the hyperparameter settings used in both stages.

Table 3 Model Training Setup and Hyperparameters

	Stage 1	Stage 2
Training Objective	Encoding Demand	Estimating Structural Parameters
Parameters	$\mathcal{M}$ (RoBERTa) and $\mathcal{H}_1$	$\mathcal{H}_2$
Targets $\vec{y}$	Bids b_2, b_3, b_4, b_5 Auction Views Auction Watchers No. Active Bidders Reserve Met Indicator $\mathbb{1}[b_1 \geq p_R]$	Bids b_2, b_3, b_4, b_5
Training Dataset	74,698 Authentic and 298,792 Synthetic Descriptions	
Validation Dataset	5,000 Authentic Descriptions	
Optimizer	AdamW	AdamW
Learning Rate	3×10^{-5} (max)	3×10^{-5} (max)
β_1	0.9	0.9
β_2	0.999	0.999
ϵ	1×10^{-8}	1×10^{-8}
Scheduler	Linear Decay w/ Warmup	Linear Decay w/ Warmup
Warmup Steps	300	1000
Training Iterations	25 epochs, batch size 64	15 epochs, batch size 32

Note: Target values in $\vec{y}$ were perturbed with a small amount of Gaussian noise during training to prevent over-fitting (aside from our indicator variable).

For the empirical analysis, only auctions that met the time-based inclusion criteria, which required that the last (largest) bids from five unique bidders be submitted within 3 hours of the auction's end time, were used in the training and validation sets. This resulted in a total of 74,698 auctions for training and 5,000 auctions withheld (i.e. not seen by the model during training) for validation analysis. More details regarding the data in provided below in Section 5.

4.1. Stage 1: Encoding Demand Embedding Vectors

The objective of Stage 1 is to train $\mathcal{M}$ to encode demand-relevant outcomes into the embeddings generated from the auction items' descriptions. Given that RoBERTa processes token sequences with a maximum context length of $\bar{\mathsf{d}} = 510$ tokens, we account for the constraint that the average auction description in our dataset contains 622 tokens (ranging from 419 at the 5^{th} percentile to 872 at the 95^{th} percentile). Since our analysis indicates that the most economically valuable information is concentrated at the beginning of the text, we do not expect this truncation to significantly impact the model's predictive capacity.

To project the demand embedding vectors $\vec{e}$ from $\mathcal{M}$ into the auction outcome space, we append a fully connected neural network layer, $\mathcal{H}_1$. This layer serves as a decoding mechanism that translates the high-dimensional text embeddings into a structured set of predictions $\hat{\vec{y}}$ for the market target outcomes $\vec{y}$. Specifically, $\mathcal{H}_1$ produces estimates for the second through fifth highest bid values (b_2, b_3, b_4, b_5), as well as auction-level engagement metrics such as total views, watchers, and unique bidders. During training, a small amount of Gaussian noise was added to the true target values of these variables to prevent over-fitting of the model. Additionally, it predicts whether the auction's reserve price p_R was met, denoted by the binary indicator $\mathbb{1}[b_1 \geq p_R]$. Given the binary nature of this target outcome, we apply a sigmoid transformation to the prediction made by $\mathcal{H}_1$ to ensure its values are in the unit interval, $[0, 1]$. Stage 1 is trained for 25 epochs using the AdamW optimizer with a maximum learning rate of 3×10^{-5} and a linear warmup decay schedule. We employ a batch size of 64 and monitor the mean squared error (MSE) loss in equation (2) across all predicted outcomes in the training dataset.

4.2. Stage 2: Estimating Structural Demand Parameters

The second stage of training refines the structured representations of demand learned in Stage 1 by estimating the latent distributions of consumer valuations and market size in our structural model $\mathcal{F}$. While Stage 1 provides direct predictions of auction outcomes, Stage 2 introduces a structural estimation framework that parameterizes the probability density functions of the demand primitives as defined in equations (3) and (4), allowing for a deeper economic interpretation of how textual descriptions influence bidder behavior.

In Stage 2, the demand embeddings $\vec{e}$ from $\mathcal{M}^*$ are frozen and passed through a nonlinear transformation layer in $\mathcal{H}_2$. This transformation extracts a lower-dimensional representation optimized for structural inference. The transformation consists of two fully connected layers with hidden dimensions of 32 and 16, each followed by LayerNorm regularization and a SiLU activation function, which introduces smooth non-linearity to enhance feature extraction while maintaining numerical stability. Separate output heads within $\mathcal{H}_2$ then estimate the parameters of the valuation distribution $F_V(v)$ and market size distribution $F_N(n)$. The optimization objective stated in equation (7) for Stage 2 minimizes the discrepancy between the $2^{\text{nd}} - 5^{\text{th}}$-largest bids submitted by unique bidders and those generated from the estimated joint order statistic distributions in equation (5). As in the Stage 1 estimation, a small amount of Gaussian noise was added to these variables during training.

Similar to Liu and Bagh (2020), we use Sobol sequences — low-discrepancy quasi-Monte Carlo (QMC) sequences — to efficiently approximate the expectations of these order statistics. This approach provides more uniform coverage of the probability space compared to standard Monte Carlo sampling (Morokoff and Caflisch 1995, Jäckel 2002), and also allows for a more stable and sample-efficient numerical integration when computing bid expectations over the estimated valuation distribution. Stage 2 is trained using the AdamW optimizer with a learning rate of 3×10^{-5} and a batch size of 32. The training process spans 15 epochs and incorporates a linear warmup schedule with 1,000 warmup steps.

Once trained, the full model consists of $\mathcal{M}^*$ and $\mathcal{H}_2^*$, which are evaluated on held-out auction data to assess generalization. The combination of Stage 1 demand encoding and Stage 2 structural framework allows us to move beyond direct outcome prediction, allowing a more interpretable framework to analyze bidder willingness to pay and auction participation. In the empirical section that follows, we validate the predictive accuracy of the model and perform counterfactual simulations to examine how different vehicle descriptions influence estimated demand.

5. Online Vehicle Auction Data – BringATrailer.com

For the empirical application of our estimation method, we collected data from bringatrailer.com (BaT), one of the largest online vehicle auction marketplaces in USA. BaT utilizes an ascending English auction mechanism with a secret reserve price p_R for their vehicle auctions. The description of each auction vehicle is written by BaT staff following a stylized template that reviews the history of the vehicle, its current condition, any modifications or restorations, and its provenance; see Appendix B for an example description to an auction listing for a 2018 Porsche 911.

Each auction begins at $1 (all prices in USD), typically with a seven day clock that eventually ends the auction. To avoid snipe bidding, BaT automatically extends the auction's clock by two minutes whenever a new bid is submitted within this threshold. When the auction concludes, the highest bid b_1 is privately compared to the vehicle's secret reserve price p_R, and automatically announced as to whether the reserve was met. If the seller's reserve price is met ($b_1^* \geq p_R$), then the winning bidder pays the seller the amount b_1^*. In addition, the winning bidder is charged a BaT transaction fee of 5% of b_1^* or $250, whichever was greater, but no more than $5,000, i.e. $\min(\max(.05 \times b_1^*, \$250), \$5,000)$. If the seller's reserve price is not met ($b_1^* < p_R$), then BaT connects the highest bidder with the seller to attempt a privately negotiated a sale price. However, regardless of the outcome, the seller is charged an unrecoverable listing fee of $99.

Important to this study, all text-based communications and bids submitted during a given auction were publicly recorded in a chat section at the bottom of the auction's webpage. While presumably designed to enhance the social aspects of the auction site, and reduce some information asymmetries, our estimation methods benefit from having every submitted bid timestamped and user-specified by BaT. We sourced all market data from the information in these chat sections.

5.1. Descriptive Summary Statistics

From July 2014 through December 2023, BaT held a total of 89,080 auctions, however only 79,698 of them met our inclusion criteria that the final bids from five unique bidders be submitted within 3 hours of the auction's end time. This decreased the likelihood that some high value bidders were not present toward the end of the auction. Additionally, recall that to mitigate snipe biding, BaT extended the end time by two minutes whenever a bid was submitted within the last two minutes. We summarize the characteristics of the auctions meeting the inclusion criteria in Table 4.

Table 4 Descriptive Statistics of sample BringATrailer.com data (July 2014 - December 2023)

79,698 Auctions	**Mean**	**St. Dev.**	**Min.**	**Max.**
Bidding Duration (Days)†	6.902	.920	.0097	21.054
$\mathbb{1}$[Sold by Dealer]	.316	.465	0	1
$\mathbb{1}[\max(b) \geq r]$ (Reserve Met)	.772	.419	0	1
No. of Active Bidders	12.007	4.522	1	39
Auction Views (1000s)	14.209	9.782	.698	358.191
Auction Watchers (1000s)	.726	.349	.060	6.105
Nominal Bid Values				
Winning Bid (in 1000s)	\$44.452	\$79.284	\$.105	\$5,360
Normalized Bid Values				
1st Largest (Winning) Bid	1	n/a	1	1
2nd Largest Bid	.969	.056	.000	.999
3rd Largest Bid From...				
...All Bidders	.941	.083	.000	.998
...Non-winning Bidders	.878	.120	.000	.998
All Bids†	.662	.261		
All Unique Bidders' Max. Bids†	.615	.295		
Fraction of Bids Submitted in Final...	**5 Min.**	**15 Min.**	**1 Hr.**	**24 Hr.**
1st Largest (Winning) Bid	1	1	1	1
2nd Largest Bid	.810	.854	.901	.966
3rd Largest Bid From...				
...All Bidders	.738	.794	.849	.935
...Non-winning Bidders	.835	.860	.891	.948
All Bids†	.160	.342	.470	.633
All Unique Bidders' Max. Bids†	.190	.270	.358	.517

199,468 Unique Bidders	**Mean**	**St. Dev.**	**Min.**	**Max.**
No. Auctions Participated	4.797	15.655	1	1856
Fraction of Bidders Who Won...Auctions	**0**	**≥1**	**≥2**	**≥5**
	.749	.251	.059	.009

† Calculated for auctions that had at least three bidders.

Approximately 31.6% of the auctions were managed by dealers, and the reserve prices were met in 77.2% of cases. Bidder participation averaged 12.007 active bidders per auction, with a minimum of 1 and maximum 39 unique bidders. For bid values, the average winning bid across the auctions was \$44,452, with the highest recorded bid reaching \$5,360,000 and the lowest at \$105. When normalized for analysis with the highest bid set to 100%, the second and third highest bids averaging 96.9% and 94.1%, respectively, of the winning bid. Notably, the third highest bids submitted by non-winning participants (i.e. those who did not ultimately submit the single largest bid) averaged slightly lower, at 87.8% of the highest bid.

The timing of bids demonstrates regular late bidding, with a significant portion of activity occurring close to the auction's conclusion. Specifically, 16% of all bids were submitted in the final five minutes, 34.2% in the last 15 minutes, 47% within the last hour, and 63.3% on the final day. While partly a consequence due to BaT default listing of auctions in ascending order of remaining time, i.e. auctions closest to end are listed first, it also reflects that it is only the most committed bidders that actively participate. Thus providing further support to that bidders are bidding up to their valuation.

The auction platform engaged a total 199,468 unique bidders during our data collection period, each participating in an average of 4.797 auctions. Despite this active engagement, 74.9% of bidders never won an auction, and only 25.1% won at least once. Furthermore, only a small fraction of the participants achieved multiple wins: 5.9% won at least twice, and a mere 0.9% won five or more auctions.

5.2. Authentic and Synthetic Data

To enhance the robustness of our estimation process, we supplement the auction house's standardized vehicle descriptions with a sample of 298,792 synthetic descriptions generated by the LLAMA 3.1 70B model (Dubey et al. 2024). This augmentation is particularly valuable in Stage 1, where the language model learns to encode descriptions into structured demand representations. The structured nature of BaT's descriptions, while useful for consistency, may limit the language model's ability to generalize across different buyer perspectives and listing styles. Synthetic descriptions mitigate this risk by introducing diverse linguistic and conceptual framing, allowing the model to better capture the heterogeneity in how vehicle attributes are perceived by bidders.[2]

To generate synthetic descriptions, we simulate the personas of four distinct automotive experts: an Independent Vehicle History Report Provider; a Certified Automotive Mechanic; a Car Enthusiast Club President; and a Vehicle Appraiser. Each expert role is designed to emphasize different aspects of the vehicle, including provenance, mechanical integrity, enthusiast appeal, and market valuation. The descriptions are generated iteratively, with each expert receiving a prompt instructing them to evaluate a vehicle based on an authentic description sourced from our dataset. This process ensures that the generated content goes beyond simple paraphrasing, embedding distinct evaluative insights that would otherwise be absent from the auction house's standardized listings. A truncated comparison of authentic and synthetic descriptions for one vehicle is presented in Table 5.

By using multiple expert perspectives, we ensure that key factors influencing auction outcomes—such as mechanical reliability, market desirability, and historical significance—are explicitly addressed in the training data. For example, while a mechanic might highlight potential repair costs and long-term reliability, an appraiser would focus on rarity, market trends, and comparative sales

[2] For other results on using diversified personas during model training, see Jandaghi et al. (2023) and Ge et al. (2024).

Table 5 Sample Description Format - Authentic vs. Synthetic

Description Format, Authentic	Description Format, Synthetic
Title: Modified 1974 BMW 3.0CS Date: February 27, 2024 Description: This 1974 BMW 3.0CS was refurbished before being acquired by the current owner in November 2021, and work included rust remediation, a repaint, an interior refresh, and updates to the brake and suspension systems...	Title: Modified 1974 BMW 3.0CS Date: February 27, 2024 Review: This 1974 BMW 3.0CS appears to be a well-maintained and tastefully modified example of a classic car. The extensive refurbishment indicates a thorough and high-quality restoration. The modifications, such as the CSL-style...

data. This methodological design not only enhances the language model's ability to extract meaningful demand signals but also improves generalization, particularly for vehicles with limited auction history or those falling outside the most commonly listed categories.

Beyond improving model flexibility, the use of synthetic data allows us to assess robustness by testing whether valuation predictions remain stable across different textual framings of the same vehicle. By systematically varying the presentation of information while maintaining factual consistency, we evaluate whether the model's demand estimates are sensitive to linguistic and contextual differences. Additionally, the presence of diverse description styles helps prevent overfitting to a single mode of vehicle presentation, reducing biases that might emerge if the model were trained exclusively on auction house-generated descriptions.

6. Empirical Analysis of Online Vehicle Auction Data

We begin our analysis of the trained model by evaluating its predictive accuracy on held-out auction data and assessing its ability to capture key economic relationships in the online vehicle auction market. Specifically, we test how well the model's Stage 1 predictions align with observed auction outcomes, including bid values, auction engagement metrics, and reserve price attainment. We then examine the structural estimates produced in Stage 2, validating the recovered valuation and market size distributions against empirical bid distributions. Finally, we conduct counterfactual simulations to analyze how variations in vehicle descriptions influence estimated demand primitives, allowing us to quantify the economic impact of textual attributes on bidding behavior and price formation.

6.1. Stage 1 Validation Results of Demand Embedding Vectors

Figure 4 presents the distribution of prediction errors on the Stage 1 targets, evaluated on the held-out validation dataset of 5,000 authentic auction descriptions. The panels representing the log bid estimates are presented in scatter plots (x-axis is target, y-axis is prediction), while the number of auction views, watchers, and unique bidders are presented as error histograms in percentage terms.

We present the nominal error of the indicator for whether the reserve price was met at the conclusion of the auction. The distributions of all bid-related errors are approximately centered around zero with an average absolute error of 15-16%, indicating that the model does not exhibit systematic bias in over- or under-predicting bid values. Similarly, the prediction errors for views and watchers show a symmetric distribution, suggesting a balanced estimation of auction engagement. The unique bidders panel exhibits a slight right skew, indicating that the model occasionally underpredicts participation in high-competition auctions.

The final panel reports the model's prediction accuracy for whether the reserve price was met. Approximately 77% of auctions in the validation dataset resulted in a sale, and the model correctly classifies roughly 75% of these outcomes. This suggests that the model has not learned much about how the auction house sets its reserve prices, likely because this process is determined by seller preferences rather than directly reflecting market demand. However, since reserve price setting is not a direct indicator of bidder valuation, this limitation has minimal implications for the model's core objective of estimating demand. Overall, the validation results indicate that the model provides a well-calibrated representation of auction outcomes, with minimal systemic biases in its predictions.

To demonstrate where the model assigns its estimated value within the description of a vehicle, we present a randomly selected description in Figures 5 and 6. Here, we provide token-level annotations that visualize the impact each token in the vehicle description has on the auction bid value submitted by the price-setting bidder (i.e. the bidder with second largest bid). Each token within the description is enclosed in a grayscale box where the depth of the gray shade represents the estimated USD value that the token adds to the bid. The darker the shade, the higher the contribution of that token to the bid value, as indicated in the gradient legend. This technique is useful for identifying significant features and attributes that influence bidding behavior.

The token-level analysis of price-setting bids reveals clear patterns in how different aspects of a vehicle's description contribute to its estimated auction price. Tokens with the highest contributions generally relate to the vehicle's identity, exclusivity, and auction-specific signals. Notably, references to the make/model, its no reserve status, and its "Manhattan Automobile Company" origin significantly increased the estimated price, suggesting that bidders value brand prestige, high-end trim levels, and a transparent auction process. Mentions of the vehicle's history in New York and Florida also contributed strongly, likely due to market desirability and reduced concerns about winter-related wear. Performance and luxury features, such as the 5.0-liter V8 engine, ventilated and massaging seats, navigation system, and 20-inch alloy wheels, were also strong price indicators, reinforcing the premium placed on high-performance specifications and rare comfort features. Conversely, lower-value contributions were associated with standard equipment, maintenance records, and general descriptors, indicating that routine service history and common vehicle features do not significantly impact bidding behavior.

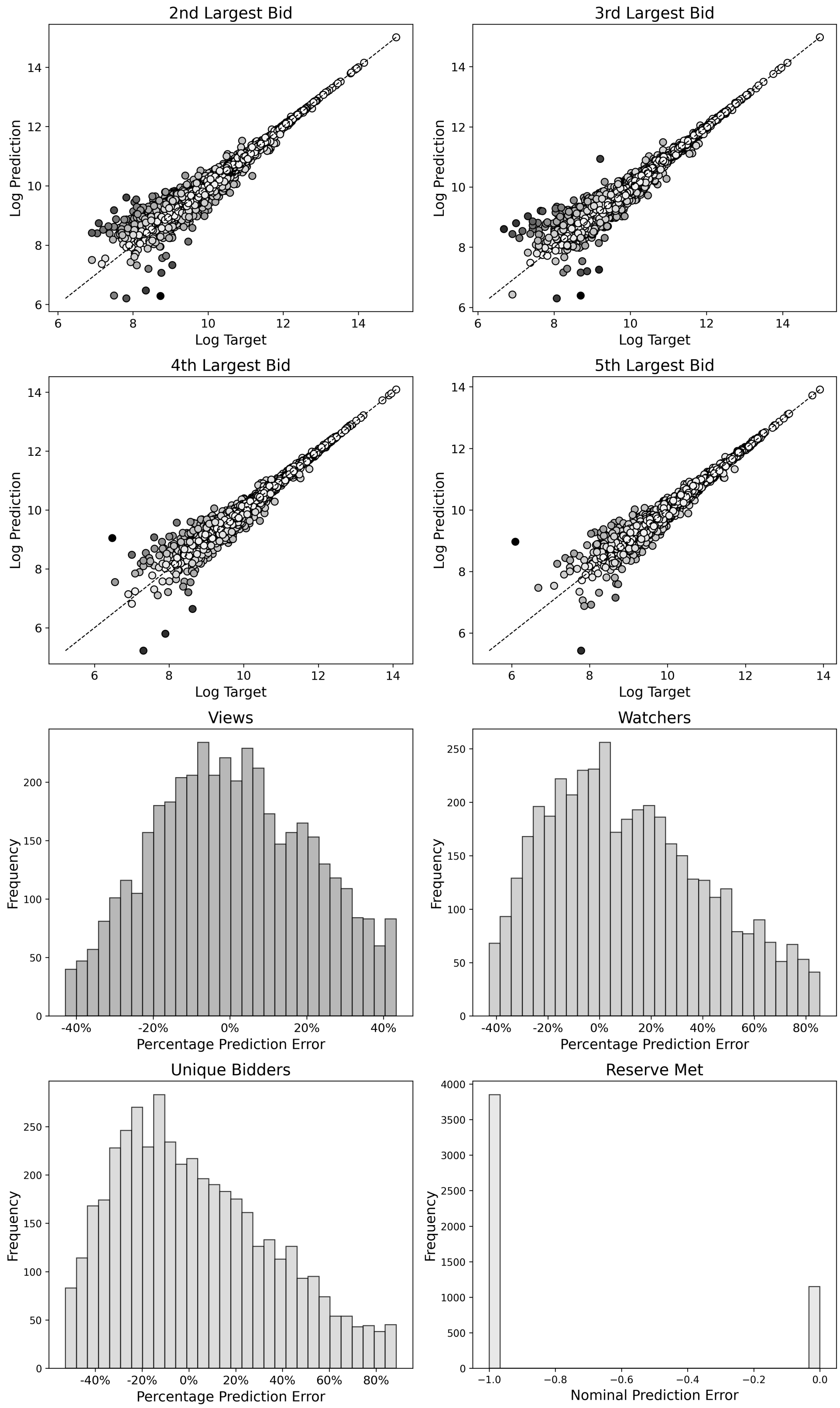
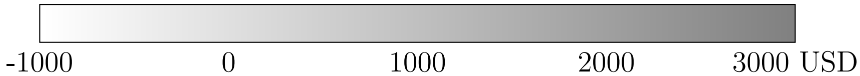


Figure 4 Stage 1 Validation Prediction Errors (sample size: 5,000)



Title: No Reserve: 2013 Jaguar XJL Portfolio

Date: June 03, 2022 at 05:45 PM

Description: This 2013 Jaguar XJL Portfolio was sold new by Manhattan Automobile Company and spent time in New York and Florida prior to the selling dealer's acquisition in 2022. The car is finished in Carnelian Red over London Tan leather upholstery and powered by a 5.0-liter V8 paired with an eight-speed automatic transmission. Equipment includes 20 inch wheels, rear tray tables, a heated steering wheel, four-zone automatic climate control, navigation, a panoramic sunroof, a Meridian audio system, and heated and ventilated seats with front seat massaging elements. The water pump was replaced in 2018 and service in preparation for the sale reportedly consisted of performing an oil change in addition to replacing the battery and the front control arms. This XJL has 42k miles and is now offered at no reserve by the selling dealer in Illinois with a clean Carfax report and a clean Pennsylvania title. The body is finished in Carnelian Red and equipped with brightwork consisting of a Leaper hood ornament, a mesh grille, side mirror caps, and window trim. Additional features include a panoramic sunroof, bi-xenon headlights with washers, vented fenders, and front and rear parking sensors. The selling dealer notes a stone chip on the hood and blemishes are visible on the bumper covers. Silver-finished 20 inch Kasuga alloy wheels wear 245/4o front and 275/35 rear Hankook Ventus V12 EVO2 tires. Stopping power is provided by four wheel disc brakes, and the car is equipped with an Adaptive Dynamics suspension. The front control arms have been replaced in preparation for the sale, according to the selling dealer. The heated and ventilated front and rear seats are trimmed in London Tan leather with Navy piping. Burl walnut veneers accent the dashboard, steering wheel, door panels, and rear tray tables, and additional amenities include front seat massage and memory settings, four-zone automatic climate control, rear sunshades, an 8 inch touchscreen infotainment system with navigation, and a 14-speaker Meridian audio system. A heated multifunction steering wheel frames a digital instrument panel that includes a 170-mph speedometer, a tachometer, and a combination gauge.

Figure 5 Stage 1 Estimation: Token's Contribution to Price-setting Bid

The factors influencing auction views differ significantly from those driving the price-setting bid, highlighting distinct mechanisms underlying bidder engagement and price determination. Tokens contributing to higher auction views tend to emphasize broad visibility and desirability, with strong impacts from the vehicle's brand identity, color, trim, and key luxury features. Mentions of Jaguar

XJL, No Reserve, and Carnelian Red substantially increased auction traffic, likely because these elements enhance listing appeal and encourage prospective bidders to explore the auction. Additionally, auction-specific timing and dealer history played a crucial role, as seen in the high contributions from auction date and time, Manhattan Automobile Company, and New York, indicating that search behavior and perceived vehicle provenance influence initial engagement.

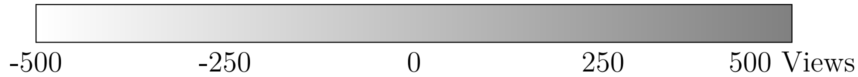
In contrast, the price-setting bid was more directly influenced by factors tied to valuation, competitive bidding, and price-setting characteristics. While brand identity and luxury cues remained important, their impact on final bid levels was more nuanced. Performance-related specifications, such as the 5.0-liter V8 engine and eight-speed automatic transmission, played a stronger role in determining bid increments than in attracting views, reflecting their importance in bidder willingness to pay. Similarly, maintenance history and vehicle condition, which had only a moderate effect on views, contributed more significantly to the price-setting bid, suggesting that bidders scrutinize reliability and upkeep more closely when committing to higher bids rather than when initially engaging with the auction.

Another key contrast is the role of unique or high-end features. Elements such as ventilated seats, panoramic sunroof, and Meridian audio system were strong drivers of views, but their impact on price-setting bids was less pronounced, indicating that while these features attract attention, they do not necessarily translate into higher willingness to pay. Conversely, mechanical integrity and major service records, which had limited influence on auction views, played a larger role in bid determination, reinforcing the idea that serious bidders weigh long-term ownership costs more heavily than casual viewers. Overall, auction views are driven by immediate listing appeal and search visibility, while the price-setting bid is influenced by valuation-relevant attributes that justify higher price commitments. This distinction underscores how different aspects of a vehicle's description engage different stages of bidder decision-making, with attention-grabbing features bringing bidders to the auction and valuation-relevant factors guiding final price formation.

The significance of this model lies in its ability to pinpoint precisely where the value is being perceived at the token level within a vehicle's description. By understanding which features or attributes carry more weight in the bidding process, sellers can strategically tailor their descriptions to emphasize these valuable aspects, potentially increasing the final bid amounts. This approach ultimately enhances the richness of the dataset, providing a more nuanced understanding of the factors that drive auction dynamics.

6.2. Stage 2 Empirical Estimation of Valuation and Market Size

Figure 6.2 presents the nominal prediction errors from the structural Stage 2 model, evaluated on the validation dataset (sample size 5,000). Each panel reports the log prediction error for the 2^{nd} through



Title: No Reserve: 2013 Jaguar XJL Portfolio

Date: June 03, 2022 at 05:45 PM

Description: This 2013 Jaguar XJL Portfolio was sold new by Manhattan Automobile Company and spent time in New York and Florida prior to the selling dealer's acquisition in 2022. The car is finished in Carnelian Red over London Tan leather upholstery and powered by a 5.0-liter V8 paired with an eight-speed automatic transmission. Equipment includes 20 inch wheels, rear tray tables, a heated steering wheel, four-zone automatic climate control, navigation, a panoramic sunroof, a Meridian audio system, and heated and ventilated seats with front seat massaging elements. The water pump was replaced in 2018 and service in preparation for the sale reportedly consisted of performing an oil change in addition to replacing the battery and the front control arms. This XJL has 42k miles and is now offered at no reserve by the selling dealer in Illinois with a clean Carfax report and a clean Pennsylvania title. The body is finished in Carnelian Red and equipped with brightwork consisting of a Leaper hood ornament, a mesh grille, side mirror caps, and window trim. Additional features include a panoramic sunroof, bi-xenon headlights with washers, vented fenders, and front and rear parking sensors. The selling dealer notes a stone chip on the hood and blemishes are visible on the bumper covers. Silver-finished 20 inch Kasuga alloy wheels wear 245/4o front and 275/35 rear Hankook Ventus V12 EVO2 tires. Stopping power is provided by four wheel disc brakes, and the car is equipped with an Adaptive Dynamics suspension. The front control arms have been replaced in preparation for the sale, according to the selling dealer. The heated and ventilated front and rear seats are trimmed in London Tan leather with Navy piping. Burl walnut veneers accent the dashboard, steering wheel, door panels, and rear tray tables, and additional amenities include front seat massage and memory settings, four-zone automatic climate control, rear sunshades, an 8 inch touchscreen infotainment system with navigation, and a 14-speaker Meridian audio system. A heated multifunction steering wheel frames a digital instrument panel that includes a 170-mph speedometer, a tachometer, and a combination gauge.

Figure 6 Stage 1 Estimation: Token's Contribution to Auction Views

5^{th} highest bids, which were estimated using the inferred valuation and market size distributions. The error distribution for each bid rank is characterized by its first and second moments. For the 2^{nd} highest bid, the mean error is 0.1505 with a standard deviation of 0.3255, indicating a distribution centered slightly above zero with moderate dispersion. The 3^{rd} highest bid exhibits a mean error of

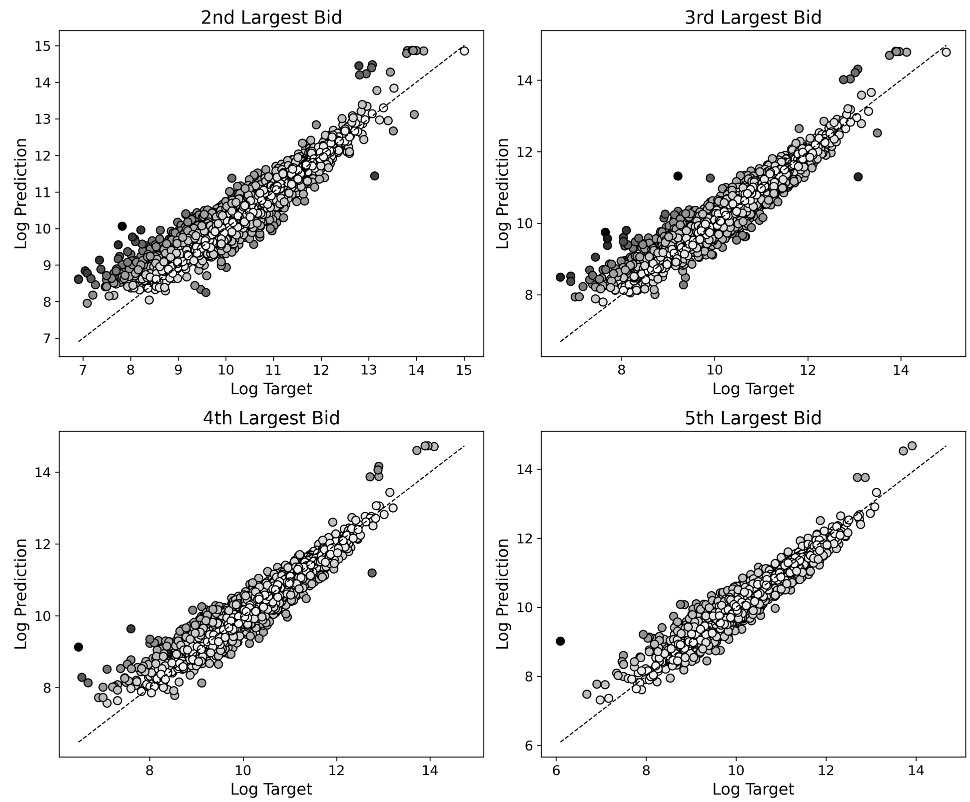


Figure 7 Validation Prediction Error on 2nd, 3rd, 4th and 5th Largest Bids (sample size: 5,000)

0.0924 and a standard deviation of 0.3092, while the 4^{th} highest bid has a mean error of 0.0637 and a standard deviation of 0.2994. The 5^{th} highest bid shows a mean error of 0.0331 with a standard deviation of 0.2866.

These estimates suggest that prediction errors are distributed with a slight positive bias, with variability largely stable across bid ranks. The standard deviation remains substantial relative to the mean error, highlighting nontrivial dispersion in prediction accuracy. While these moments provide a high-level characterization of the error structure, additional distributional analysis may be necessary to identify whether deviations arise from model misspecification or unmodeled heterogeneity in the underlying bid-generating process.

6.3. Zero-Shot Learning: The DMC DeLorean

A natural test of this model's ability to generalize beyond its training distribution is to evaluate its predictive performance on vehicles that were entirely absent from the training set. To assess this capability, we deliberately withheld a specific make and model from both the training and validation datasets: the DMC DeLorean. The vehicle, famous from the *Back to the Future* movies, possesses a distinct combination of design elements, historical significance, and cultural associations that are unlikely to be collectively shared by any other vehicles in the dataset. Consequently, accurate predictions on this category of vehicles would provide strong evidence that the model has successfully learned generalizable demand patterns within the target market.

In total, the dataset contained 366 auctions for DMC DeLoreans, none of which were seen by the model during training. Figure 6.3 presents the nominal prediction errors for these zero-shot evaluations, focusing on the model's Stage 2 predicted bids from the actual second-to-fifth largest bids observed that met our inclusion criteria in each auction. The prediction errors for the zero-shot evaluation of the DMC DeLorean exhibit distinct characteristics relative to the broader validation set. Across all bid ranks, the mean errors are negative, ranging from -0.0354 to -0.0876, suggesting a systematic tendency toward underprediction. The median errors follow a similar pattern, with values between -0.0219 and -0.0743, indicating that the central tendency of the error distribution is consistently below zero. The standard deviation of the errors remains relatively stable across bid ranks, ranging from 0.1604 to 0.1801. This suggests that while the model exhibits a slight downward bias in its predictions for the DeLorean, the overall dispersion of errors is comparable to that observed in the general validation dataset. The root mean squared error (RMSE) values, which vary between 0.1824 and 0.1860, further reinforce this observation, showing little fluctuation across bid ranks. One possible explanation for the systematic underprediction is the model's inability to fully account for the cultural and historical significance of the DMC DeLorean, particularly its association with the *Back to the Future* franchise. Unlike standard vehicle attributes such as performance, production volume, or brand reputation, the DeLorean's demand dynamics are heavily influenced by its status as an iconic pop culture artifact. This historical and cultural context is difficult to infer from structured auction data alone, as the model lacks direct access to external sources that encode the persistent influence of media-driven valuation. Consequently, the model may fail to recognize that bidders assign additional value to the DeLorean not merely as a collectible vehicle, but as a symbol of 1980s nostalgia and science fiction fandom.

Our fine-tuned model's ability to make robust out-of-sample predictions on vehicles it has never encountered directly stems from the deep representations learned during pretraining (Wei et al. 2021). Large-scale language models, such as RoBERTa, are trained on extensive corpora that include information about a wide range of consumer goods, historical events, and market trends. It is likely that the foundational model trained in Liu et al. (2019) was exposed to content about the DeLorean, including its production history, collector status, and pop culture significance. During the fine-tuning process in Stage 1, these latent representations were likely preserved and refined within the context of auction markets, allowing the model to infer a reasonable valuation for the DeLorean by positioning it relative to other vehicles with similar attributes in the dataset. However, this explanation remains speculative, as testing whether and how these representations were preserved is beyond the scope of this paper. Future research could more directly investigate the extent to which pretrained

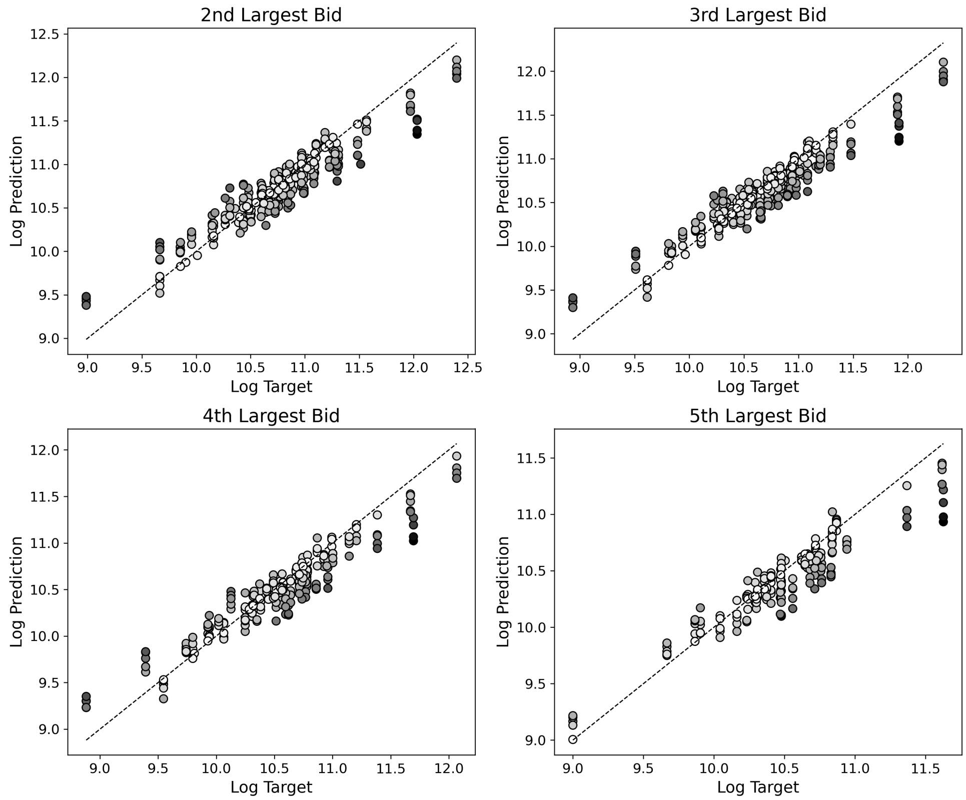


Figure 8 **Zero-shot Predictions: DMC DeLorean (sample size: 366)**

language models retain and adapt knowledge representations during fine-tuning, particularly in economic applications (see Templeton 2024).

6.4. Counterfactual Simulations: Systematically Varying Vehicle Mileage

To assess the economic consistency of our trained model, we conduct a series of counterfactual simulations that systematically vary a key determinant of vehicle value: mileage. Vehicle mileage is one of the most robust indicators of depreciation, influencing resale values across all makes and models. Therefore, as a robustness check, we test whether the model's price-setting bid predictions exhibit an economically reasonable response to changes in reported mileage. Due to the computational cost of this analysis, we randomly select a subset of 1,000 vehicles from the validation dataset and use a large language model (LLAMA 3.1 70B) to generate modified versions of each vehicle's description while holding all other details constant. Specifically, we replace the mileage value in the original description with one of six predetermined values: 25k, 50k, 75k, 100k, 125k, and 150k miles. Each edited description is then processed by the trained Stage 2 model to generate a new set of predicted price-setting bids.

Figure 9 presents the results of this experiment, where the price-setting bid predictions are normalized to 1 for the 25k-mile counterfactual and 0 for the 150k-mile counterfactual. Across all vehicles

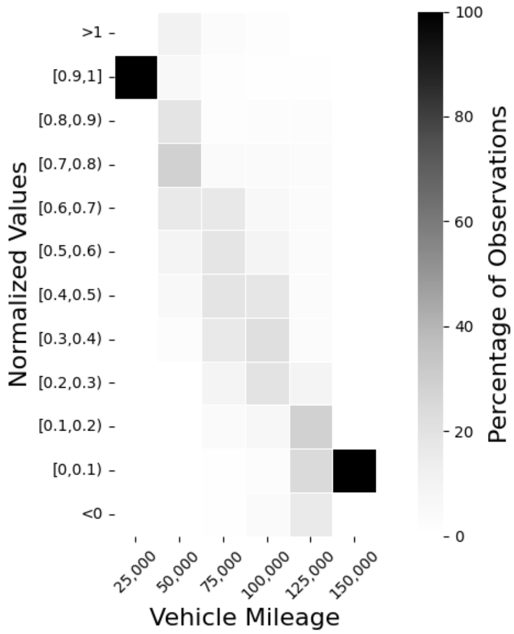


Figure 9 **Counterfactual Mileage Simulations (sample size: 1,000)**

included in the analysis, the model predicts a monotonic decline in the price-setting bid as mileage increases. This finding aligns with standard economic expectations regarding vehicle depreciation, and reinforcing the model's ability to appropriately internalize the role of accumulated mileage in determining vehicle value.

These results also highlight the utility of our framework for evaluating the marginal effects of key vehicle attributes on expected auction outcomes. By systematically modifying vehicle descriptions and observing the model's predictions, researchers and practitioners can conduct detailed counterfactual analyses that would be infeasible to observe in real-world auction data. Future work could extend this approach by examining other vehicle characteristics, such as accident history, number of previous owners, or specific aftermarket modifications, to further explore the model's sensitivity to economically relevant factors.

7. Conclusion

This paper introduces a novel approach to demand estimation that integrates language models with structural econometric techniques to infer primitives of demand from unstructured textual descriptions. By training a RoBERTa-based model in a two-stage process, we demonstrate how natural language descriptions of auctioned vehicles can be systematically transformed into structured economic representations that capture latent demand primitives. Our empirical validation using online vehicle auction data confirms that the proposed model effectively encodes market-relevant information, yielding competitive predictive accuracy on auction outcomes while preserving interpretability through its structural estimation framework.

A key strength of our approach is its ability to generalize beyond the training distribution, as evidenced by the model's performance on zero-shot predictions for the DMC DeLorean. This result highlights the potential of language models to leverage their pre-trained knowledge when making economic inferences in unfamiliar contexts, a property that could be further explored in other market settings. Additionally, our counterfactual simulations on vehicle mileage confirm the model's ability to produce economically coherent predictions, reinforcing its capacity to extract meaningful demand signals from textual descriptions.

These findings have several important implications for both researchers and practitioners. For firms and auction platforms, the ability to extract consumer demand signals directly from product descriptions offers a new avenue for pricing strategy and inventory management. Rather than relying solely on historical sales data, firms can leverage language models to predict demand for new products, optimize product descriptions for maximum market appeal, and conduct counterfactual analysis to explore pricing and positioning strategies. From a methodological perspective, our study contributes to the growing literature on machine learning applications in economics by demonstrating how deep learning models can be integrated with structural estimation techniques to improve demand inference.

Despite these contributions, important avenues for future research remain. First, while our approach successfully captures demand primitives from textual data, further work is needed to understand how fine-tuned language models retain and adapt pre-trained knowledge for economic inference. The extent to which language models internalize and preserve economic reasoning during domain-specific training is an open question that warrants further exploration using interpretability methods, such as feature attribution and causal probing techniques. Second, while our empirical application focuses on English auctions, the underlying framework could be extended to other market settings, including posted-price markets, bargaining environments, and two-sided platforms where product descriptions and buyer preferences interact dynamically. Lastly, given the rapid evolution of language model architectures, future studies could investigate how different model families (e.g., multimodal models with vision capabilities) compare in their ability to encode and generalize economic information.

References

Aceves, Pedro, and James A Evans. 2024. "Mobilizing conceptual spaces: How word embedding models can inform measurement and theory within organization science." *Organization Science*, 35(3): 788–814.

Adam, Hammaad, Pu He, and Fanyin Zheng. 2024. "Machine learning for demand estimation in long tail markets." *Management Science*, 70(8): 5040–5065.

Ash, Elliott, and Stephen Hansen. 2023. "Text algorithms in economics." *Annual Review of Economics*, 15: 659–688.

Athey, Susan. 2018. "The impact of machine learning on economics." In *The economics of artificial intelligence: An agenda.* 507–547. University of Chicago Press.

Athey, Susan, and Guido W Imbens. 2019. "Machine learning methods that economists should know about." *Annual Review of Economics*, 11(1): 685–725.

Avery, Christopher. 1998. "Strategic jump bidding in English auctions." *The Review of Economic Studies*, 65(2): 185–210.

Berger, Jonah, Ashlee Humphreys, Stefan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel. 2020. "Uniting the tribes: Using text for marketing insight." *Journal of Marketing*, 84(1): 1–25.

Chen, Jieneng, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L Yuille, and Yuyin Zhou. 2021. "Transunet: Transformers make strong encoders for medical image segmentation." *arXiv preprint arXiv:2102.04306.*

Chen, Zenan, and Jason Chan. 2024. "Large Language Model in Creative Work: The Role of Collaboration Modality and User Expertise." *Management Science*, 70(12): 3450–3472.

Cropper, Maureen L, Leland B Deck, and Kenenth E McConnell. 1988. "On the choice of funtional form for hedonic price functions." *Review of Economics and Statistics*, 668–675.

de Kok, Ties. 2025. "ChatGPT for Textual Analysis? How to Use Generative LLMs in Accounting Research." *Management Science*, 71(1): 123–145.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805.*

Dillion, Danica, Niket Tandon, Yuling Gu, and Kurt Gray. 2023. "Can AI language models replace human participants?" *Trends in Cognitive Sciences*, 27(7): 597–600.

Dubey, Abhimanyu, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. 2024. "The llama 3 herd of models." *arXiv preprint arXiv:2407.21783.*

Easley, Robert F, and Rafael Tenorio. 2004. "Jump bidding strategies in internet auctions." *Management Science*, 50(10): 1407–1419.

Elhage, Nelson, Tristan Hume, Catherine Olsson, Nicholas Schiefer, Tom Henighan, Shauna Kravec, Zac Hatfield-Dodds, Robert Lasenby, Dawn Drain, Carol Chen, et al. 2022. "Toy models of superposition." *arXiv preprint arXiv:2209.10652*.

Farrell, Max H, Tengyuan Liang, and Sanjog Misra. 2020. "Deep learning for individual heterogeneity: An automatic inference framework." *arXiv preprint arXiv:2010.14694*.

Fenton, Victor M, and A Ronald Gallant. 1996*a*. "Convergence rates of SNP density estimators." *Econometrica*, 719–727.

Fenton, Victor M, and A Ronald Gallant. 1996*b*. "Qualitative and asymptotic performance of SNP density estimators." *Journal of Econometrics*, 74(1): 77–118.

Gallant, A Ronald, and Douglas W Nychka. 1987. "Semi-nonparametric maximum likelihood estimation." *Econometrica*, 363–390.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as data." *Journal of Economic Literature*, 57(3): 535–574.

Ge, Tao, Xin Chan, Xiaoyang Wang, Dian Yu, Haitao Mi, and Dong Yu. 2024. "Scaling synthetic data creation with 1,000,000,000 personas." *arXiv preprint arXiv:2406.20094*.

Goli, Ali, and Amandeep Singh. 2024. "Frontiers: Can Large Language Models Capture Human Preferences?" *Marketing Science*, 43(4): 709–722.

Hornik, Kurt, Maxwell Stinchcombe, and Halbert White. 1989. "Multilayer feedforward networks are universal approximators." *Neural Networks*, 2(5): 359–366.

Iskhakov, Fedor, John Rust, and Bertel Schjerning. 2020. "Machine learning and structural econometrics: contrasts and synergies." *The Econometrics Journal*, 23(3): S81–S124.

Jäckel, Peter. 2002. *Monte Carlo methods in finance.* Vol. 5, John Wiley & Sons.

Jandaghi, Pegah, XiangHai Sheng, Xinyi Bai, Jay Pujara, and Hakim Sidahmed. 2023. "Faithful persona-based conversational dataset generation with large language models." *arXiv preprint arXiv:2312.10007*.

Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. 2021. "Highly accurate protein structure prediction with AlphaFold." *Nature*, 596(7873): 583–589.

Li, Bohan, Hao Zhou, Junxian He, Mingxuan Wang, Yiming Yang, and Lei Li. 2020. "On the sentence embeddings from pre-trained language models." *arXiv preprint arXiv:2011.05864*.

Lin, Fangyu, Sagar Samtani, Hongyi Zhu, Brandimarte Laura, and Hsinchun Chen. 2024. "Automated Analysis of Changes in Privacy Policies: A Structured Self-Attentive Sentence Embedding Approach." *MIS Quarterly*, 48(4).

Liu, De, and Adib Bagh. 2020. "Preserving bidder privacy in assignment auctions: design and measurement." *Management Science*, 66(7): 3162–3182.

Liu, Xiao. 2023. "Deep learning in marketing: a review and research agenda." *Artificial Intelligence in Marketing*, 239–271.

Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. "Roberta: A robustly optimized bert pretraining approach." *arXiv preprint arXiv:1907.11692*.

Ludwig, Jens, and Sendhil Mullainathan. 2024. "Machine learning as a tool for hypothesis generation." *Quarterly Journal of Economics*, 139(2): 751–827.

Lundberg, Scott M, and Su-In Lee. 2017. "A Unified Approach to Interpreting Model Predictions." In *Advances in Neural Information Processing Systems 30.* , ed. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, 4765–4774. Curran Associates, Inc.

Lu, Yulong, and Jianfeng Lu. 2020. "A universal approximation theorem of deep neural networks for expressing probability distributions." *Advances in Neural Information Processing Systems*, 33: 3094–3105.

Ma, Liye, and Lan Luo. 2023. "Beyond Fake or Genuine–The Effect of Large Language Models (LLMs) on the Content and Sentiment of Product Reviews." *USC Marshall School of Business Research Paper Sponsored by iORB, No. Forthcoming*.

McCarthy, Shawn, and Gita Alaghband. 2024. "Fin-ALICE: Artificial Linguistic Intelligence Causal Econometrics." *Journal of Risk and Financial Management*, 17(12): 537.

Mikolov, Tomáš, Wen-tau Yih, and Geoffrey Zweig. 2013. "Linguistic regularities in continuous space word representations." 746–751.

Morokoff, William J, and Russel E Caflisch. 1995. "Quasi-monte carlo integration." *Journal of Computational Physics*, 122(2): 218–230.

Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives*, 31(2): 87–106.

Netzer, Oded, Arthur Lemaire, and Michal Herzenstein. 2019. "When words sweat: Identifying signals for loan default in the text of loan applications." *Journal of Marketing Research*, 56(6): 960–980.

Olah, Chris, Nick Cammarata, Ludwig Schubert, Gabriel Goh, Michael Petrov, and Shan Carter. 2020. "Zoom in: An introduction to circuits." *Distill*, 5(3): e00024–001.

Puranam, Dinesh, Vrinda Kadiyali, and Vithala Narayan. 2021. "The impact of increase in minimum wages on consumer perceptions of service: A transformer model of online restaurant reviews." *Marketing Science*, 40(5): 985–1004.

Qiu, Xipeng, Tianxiang Sun, Yige Xu, Yunfan Shao, Ning Dai, and Xuanjing Huang. 2020. "Pre-trained models for natural language processing: A survey." *Science China Technological Sciences*, 63(10): 1872–1897.

Radford, Alec, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. "Language models are unsupervised multitask learners." *OpenAI blog*, 1(8): 9.

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

Templeton, Adly. 2024. *Scaling monosemanticity: Extracting interpretable features from claude 3 sonnet.* Anthropic.

Timoshenko, Anna, and John R. Hauser. 2019. "Identifying customer needs from user-generated content." *Marketing Science*, 38(1): 1–20.

Turian, Joseph, Lev Ratinov, and Yoshua Bengio. 2010. "Word representations: a simple and general method for semi-supervised learning." 384–394.

Vafa, Keyon, Susan Athey, and David M Blei. 2024. "Estimating Wage Disparities Using Foundation Models." *arXiv preprint arXiv:2409.09894*.

Van Giffen, Benjamin, Dennis Herhausen, and Tobias Fahse. 2022. "Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods." *Journal of Business Research*, 144: 93–106.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. "Attention is all you need." *Advances in Neural Information Processing Systems*, 30.

Verma, Sahil, Varich Boonsanong, Minh Hoang, Keegan Hines, John Dickerson, and Chirag Shah. 2024. "Counterfactual explanations and algorithmic recourses for machine learning: A review." *ACM Computing Surveys*, 56(12): 1–42.

Vylomova, Ekaterina, Laura Rimell, Trevor Cohn, and Timothy Baldwin. 2015. "Take and took, gaggle and goose, book and read: Evaluating the utility of vector differences for lexical relation learning." *arXiv preprint arXiv:1509.01692*.

Wang, Mengxin, Dennis J Zhang, and Heng Zhang. 2024. "Large Language Models for Market Research: A Data-augmentation Approach." *arXiv preprint arXiv:2412.19363*.

Wang, Wen, Beibei Li, Xueming Luo, and Xiaoyi Wang. 2023. "Deep Reinforcement Learning for Sequential Targeting." *Management Science*, 69(9): 2789–2805.

Wang, Wen, Mi Zhou, Beibei Li, and Honglei Zhuang. 2025. "Predicting instructor performance in online education: An interpretable hierarchical transformer with contextual attention." *Information Systems Research*.

Wei, Jason, Maarten Bosma, Vincent Y Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M Dai, and Quoc V Le. 2021. "Finetuned language models are zero-shot learners." *arXiv preprint arXiv:2109.01652*.

Wolfstetter, Elmar. 1996. "Auctions: an introduction." *Journal of Economic Surveys*, 10(4): 367–420.

Wu, Patrick Y, Jonathan Nagler, Joshua A Tucker, and Solomon Messing. 2023. "Large language models can be used to estimate the latent positions of politicians." *arXiv preprint arXiv:2303.12057*.

Yun, Chulhee, Srinadh Bhojanapalli, Ankit Singh Rawat, Sashank J Reddi, and Sanjiv Kumar. 2019. "Are transformers universal approximators of sequence-to-sequence functions?" *arXiv preprint arXiv:1912.10077*.

Zhu, Jian-Qiao, Haijiang Yan, and Thomas L Griffiths. 2024. "Language Models Trained to do Arithmetic Predict Human Risky and Intertemporal Choice." *arXiv preprint arXiv:2405.19313*.

Appendix A: Proofs of Propositions

A.1. Proposition 1

Let d denote a text description and $\vec{y} \in \mathbb{R}^m$ represent the vector of m economic outcomes associated with d. Let $\mathcal{M}$ be a language model that maps a text description d to a set of embedding vectors $\{\vec{e}_j\} \in \mathbb{R}^q$, where $j \in \{1, 2, ..., \mathsf{d}_{\max}\}$ represents the token position. Let $\mathcal{H}_1$ be a projection function that maps these embeddings to predicted economic outcomes $\hat{\vec{y}} \in \mathbb{R}^m$. The composition of these functions gives us $\hat{\vec{y}} = \mathcal{H}_1 \circ \mathcal{M}(d)$.

To establish the claim, we must show that for any error tolerance $\epsilon > 0$, there exist parameterized versions of these models $\mathcal{M}^*$ and $\mathcal{H}_1^*$ such that

$$\|y - \mathcal{H}_1^* \circ \mathcal{M}^*(d)\|_2 < \epsilon$$

for all (d, y) pairs in our dataset D.

By the Universal Approximation Theorem for Transformers (Yun et al. 2019)[3], for any continuous function f mapping sequences to sequences and any $\epsilon > 0$, there exists a transformer network T such that

$$\|f(x) - T(x)\|_2 < \epsilon/2$$

for all inputs x in the domain. Similarly, by the classical Universal Approximation Theorem for Feed-Forward Networks (Hornik, Stinchcombe and White 1989), for any continuous function g and any $\epsilon > 0$, there exists a feed-forward network H such that

$$\|g(\vec{e}) - H(\vec{e})\|_2 < \epsilon/2$$

for all embeddings $\vec{e}$ in the relevant domain. Let f^* be an ideal function that maps text descriptions directly to embeddings that perfectly encode the economic outcomes, and g^* be a function that maps these ideal embeddings to the true economic outcomes. By the first theorem, there exists a language model $\mathcal{M}^*$ such that

$$\|f^*(d) - \mathcal{M}^*(d)\|_2 < \delta$$

where δ is chosen to ensure that the subsequent projection error is sufficiently small. By the second theorem, there exists a projection function $\mathcal{H}_1^*$ such that

$$\|g^*(\vec{e}) - \mathcal{H}_1^*(\vec{e})\|_2 < \epsilon/2$$

for all embeddings $\vec{e}$ in the range of $\mathcal{M}^*$. Assuming $\mathcal{H}_1^*$ is L-Lipschitz continuous,

$$\|\mathcal{H}_1^* \circ f^*(d) - \mathcal{H}_1^* \circ \mathcal{M}^*(d)\|_2 \leq L \cdot \|f^*(d) - \mathcal{M}^*(d)\|_2 < L \cdot \delta.$$

We can choose $\delta = \epsilon/(2L)$ to ensure $L \cdot \delta < \epsilon/2$. Combining these inequalities through the triangle inequality,

$$\begin{aligned}
\|y - \mathcal{H}_1^* \circ \mathcal{M}^*(d)\|_2 &= \|g^*(f^*(d)) - \mathcal{H}_1^* \circ \mathcal{M}^*(d)\|_2 \\
&\leq \|g^*(f^*(d)) - \mathcal{H}_1^*(f^*(d))\|_2 + \|\mathcal{H}_1^* \circ f^*(d) - \mathcal{H}_1^* \circ \mathcal{M}^*(d)\|_2 \\
&< \epsilon/2 + L \cdot \delta \\
&< \epsilon/2 + \epsilon/2 \\
&= \epsilon
\end{aligned}$$

[3] See also Lu and Lu (2020).

Therefore, the composition $\mathcal{H}_1^* \circ \mathcal{M}^*$ can approximate the mapping from text descriptions to economic outcomes with arbitrary precision ϵ. To establish that the economic information can be recovered from the embeddings, we note that $\mathcal{H}_1^*$ serves as a decoding function. The embeddings $\{\vec{e}_j\}$ produced by $\mathcal{M}_1^*$ must contain all the information needed to reconstruct y within the specified error bound, otherwise $\mathcal{H}_1^* \circ \mathcal{M}^*(d)$ could not achieve the arbitrary precision approximation we proved above.

Appendix B: Dataset Details

Example description from our dataset:

> "This 2018 Porsche 911 GT2 RS has 15k miles and is powered by a twin-turbocharged 3.8L flat-six linked with a seven-speed PDK dual-clutch automatic transaxle and a limited-slip differential. It is finished in black over black leather and Alcantara upholstery and equipment includes the $18k Weissach Package, the Sport Chrono Package, 20 inch and 21 inch magnesium center-lock wheels, carbon-ceramic brakes with yellow calipers, a front-axle lift system, rear-axle steering, LED headlights with Porsche Dynamic Light System (PDLS), a rear wing, carbon-fiber full bucket seats, Porsche Communication Management (PCM), navigation, automatic climate control, a CD stereo, a Bose sound system, the Light Design Package, and an extended-range fuel tank. The car spent time in Pennsylvania, Montana, and California through the current owner's acquisition in 2020. This 991 GT2 RS is now offered on dealer consignment with factory literature, an accident-free Carfax report, and a clean California title. Weight savings were targeted with the GT2 RS program, and Porsche fitted multiple carbon-fiber components including the hood, front fender vents, side intakes, side mirrors, and rear wing. The Weissach Package reduces weight further through the use of carbon fiber for the roof panel and the anti-roll bars. Downforce is increased with a large rear wing, underbody diffuser, and a deeper front fascia, while NACA ducts in the front trunk lid aid in brake cooling. This example is finished in black and features LED headlights with Porsche Dynamic Lighting System (PDLS), fog lights, dual exhaust outlets, and fender vents. Paint protection film has reportedly been applied to the exterior. Staggered-diameter 20 inch and 21 inch magnesium center-lock wheels are mounted with Michelin Pilot Sport Cup 2 R tires measuring 265/35 up front and 325/30 at the rear. The car is equipped with adjustable dampers and anti-roll bars, Porsche Active Suspension Management, Porsche Stability Management, and rear-axle steering. Porsche Ceramic Composite Brakes were standard on the GT2 RS and feature yellow calipers and cross-drilled rotors. The optional front-axle lift raises the car on command to help prevent scraping."