Adventures in Demand Analysis Using AI

Martin Spindler

Presentation at the 13th ECB Conference on Forecasting Techniques, Frankfurt am Main, Germany

based on joint work with
Philipp Bach (FU Berlin), Victor Chernozhukov (MIT), Sven Klaassen (UHH), Jan Teichert-Kluge (UHH), Suhas Vijaykumar (UCSD)

Motivation

- **Early JASA tradition:** Almost a century ago, *The Journal of the American Statistical Association* published foundational empirical studies of demand (Wright, Working, Schultz, Mills, Stigler), moving economics toward quantitative measurement and establishing econometrics.
- **Hedonic modeling roots:** Classical work on hedonics (Lancaster, 50s; Griliches, 1971) showed how product attributes drive prices and demand. Pat Bajari, Pakes, and others set the stage for modern empirical work in early 2000s.
- **Today's AI/ML opportunities:** We'll use AI-generated product representations from text and images to better represent the products than the standard tabular covariates.

- **Amazon application:** Using toy car sales ranks and prices on *Amazon.com*, the paper demonstrates how multimodal transformer-based models (text, images, tabular data) generate demand-relevant and price-relevant embeddings.
- **Findings & implications:** These embeddings dramatically improve predictions of price and quantity, capturing subtle features like branding, quality, and visual appeal, thereby strengthening the hedonic approach to demand.
- **Finding for Price Elasticity**. The AI embeddings are key determinants of price sensitivity/elasticity of demand. While the embeddings are strong effect modifiers, they are very weak confounders – once the past visibility signals and prices are accounted for.

Background Papers and Collaborators

- P. Bajari, Z. Cen, V. Chernozhukov, M. Manukonda, J. Wang, R. Huerta, J. Li, L. Leng, G. Monokroussos, S. Vijaykumar, S. Wan (2021): *Hedonic Prices and Quality Adjusted Price Indices Powered by AI*. CEMMAP Working Paper CWP04/21.
- P. Bach, V. Chernozhukov, S. Klaassen, J. Teichert-Kluge, M. Spindler, S. Vijaykumar (2024): *Adventures in Demand Analysis Using AI*. arXiv:2501.00382v1.
- P. Bach, V. Chernozhukov, S. Klaassen, J. Teichert-Kluge, M. Spindler, S. Vijaykumar (2024): *Estimation of Causal Effects with Multimodal Data*. arXiv:2402.01785.
- V. Chernozhukov, C. Hansen, N. Kallus, M. Spindler, V. Syrgkanis (2024): *Applied Causal Inference Powered by ML and AI*. CausalML-Book.org.

ELSEVIER

Journal of Econometrics
Volume 251, September 2025, 106052

Hedonic prices and quality adjusted price indices powered by AI ☆

P. Bajari [a 1], Z. Cen [b 1], V. Chernozhukov [c 1], M. Manukonda [d 1], S. Vijaykumar [e 1], J. Wang [f 1], R. Huerta [d 2], J. Li [d], L. Leng [g], G. Monokroussos [h 2], S. Wang

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.jeconom.2025.106052 ↗ Get rights and content ↗

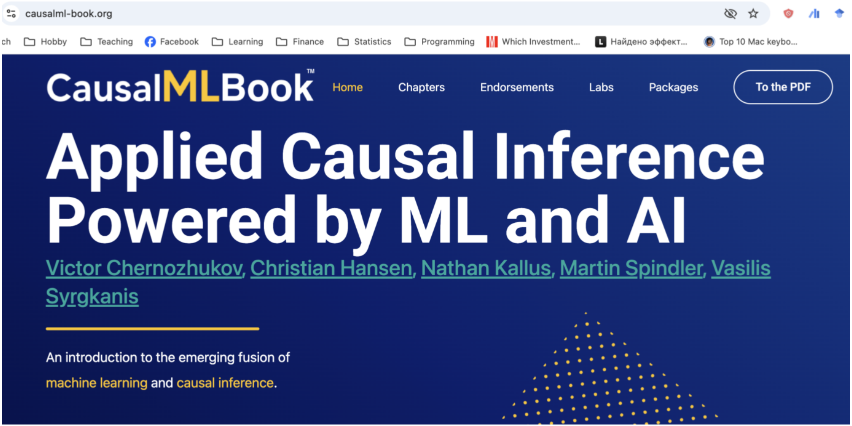
[Submitted on 31 Dec 2024]

Adventures in Demand Analysis Using AI

Philipp Bach, Victor Chernozhukov, Sven Klaassen, Martin Spindler, Jan Teichert-Kluge, Suhas Vijaykumar

This paper advances empirical demand analysis by integrating multimodal product representations derived from artificial intelligence (AI). Us
\textit{this http URL}, we combine text descriptions, images, and tabular covariates to represent each product using transformer-based emb
capture nuanced attributes, such as quality, branding, and visual characteristics, that traditional methods often struggle to summarize. More
embeddings for causal inference tasks. We show that the resulting embeddings substantially improve the predictive accuracy of sales ranks a
more credible causal estimates of price elasticity. Notably, we uncover strong heterogeneity in price elasticity driven by these product-specif
that AI-driven representations can enrich and modernize empirical demand analysis. The insights generated may also prove valuable for app

Comments: 42 pages, 9 figures
Subjects: **General Economics (econ.GN)**; Artificial Intelligence (cs.AI); Applications (stat.AP); Machine Learning (stat.ML)
Cite as: arXiv:2501.00382 **[econ.GN]**
(or arXiv:2501.00382v1 **[econ.GN]** for this version)
https://doi.org/10.48550/arXiv.2501.00382


causalml-book.org
ch
Hobby
Teaching
Facebook
Learning
Finance
Statistics
Programming
Which Investment...
Найдено эффект...
Top 10 Mac keybo...
CausalMLBook™
Home
Chapters
Endorsements
Labs
Packages
To the PDF
Applied Causal Inference Powered by ML and AI
Victor Chernozhukov, Christian Hansen, Nathan Kallus, Martin Spindler, Vasilis Syrgkanis
An introduction to the emerging fusion of machine learning and causal inference.

A Toy Example

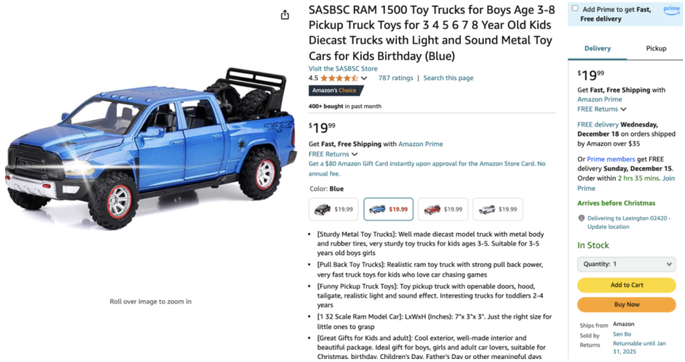


Figure: A toy example with image and text description

Toy Data for the Toy Example

- Scraped toy truck data from Keepa.com
- Sales rank of each item i at time point t
- List price of each item at each time point
- Text description of the product
- Image of the product
- Other tabular features (e.g., ratings, browse node data)

Quantity and Price Signals

- **Quantity signal:**

$$Q_{it} = \log(1/\text{Time-Averaged Sales Rank of } i \text{ in period } t)\,.$$

- **Price signal:**

$$P_{it} = \log(\text{Time-Averaged Price of } i \text{ in period } t)\,.$$

- **Data structure:** Each period $t = 1, \ldots, T$ spans 4 weeks. We set $T = 12$ periods in total, directly adjacent to each other.
- **Temporal changes:** We also examine changes in signals across periods:

$$\Delta Q_{it} := Q_{it} - Q_{i(t-1)}, \qquad \Delta P_{it} := P_{it} - P_{i(t-1)}.$$

We call the log inverse rank the "quantity signal" and log price the "price signal". Under the power law, log actual sales is **proportional** to log inverse rank.

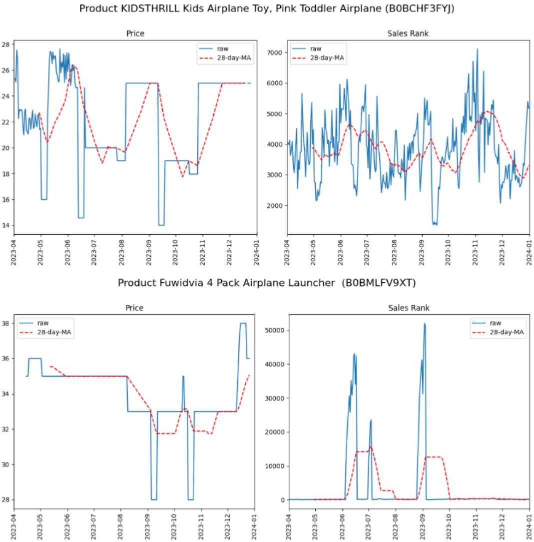


Figure 2: Price and sales rank series for two example products.

AI-Based Representation of Products

We will use various encoder models based on Transformers to convert product data into useful numerical features:

- Convert text description to dense embeddings D_{it} using Transformers (RoBERTa model; also LLama 3)
- Convert image to dense embeddings I_{it} using Transformers (BEiT model)
- Convert tabular data to embeddings T_{it} using Transformers (SAINT model)

Key Ideas: Self-Supervision, Attention, Fine-Tuning

- **Self-supervision.** Turn "unlabeled" data into massive amounts of "labeled" data by masking. For example,

$$S = (\text{Well made diecast model truck with metal body.}) \Longrightarrow$$

$$W = (\text{“Well made [m] model truck with [m] body”}, \quad Y = S$$

  Predict Y using W via neural networks. Use hidden layers of neural networks as representations E —"embeddings"—of S.
- **Attention Mechanisms**: Capture important interactions or co-occurrences of words and pieces – “Attention Is All You Need” by Vaswani, et al.
- **(Causal) Fine-Tuning:** Update embeddings to perform well on downstream tasks—for example, predicting quantity and price signals. Estimation of causal effects relies on residualization of price and quantity, and we fine-tune on the residualization tasks.

Diagram

$$
\begin{tikzcd}
 & & A_{T_i} & \\
Z_i = \begin{bmatrix} \text{Text}_i \\ \text{Image}_i \end{bmatrix} \arrow[r, mapsto, "e"] & E_i := \arrow[r, phantom, ""] & \begin{bmatrix} W_i \\ I_i \end{bmatrix} \arrow[u] \arrow[d] \arrow[r, mapsto, "m"] & \{\hat{Q}_{it}, \hat{P}_{it}\}_{t=1}^{T}. \\
 & & A_{I_i} &
\end{tikzcd}
$$

- The neural embedding map e maps text and images Z into numerical vectors E that are used to predict "auxiliary" targets—masked chunks of text and images.
- The neural map m takes the embeddings and predicts labels in downstream tasks related to causal analysis. This part can be used for fine-tuning e – updating e to perform better at the downstream task.

Embeddings: AI Representation of Products

We concatenate the embeddings: $E_{it} = (D_{it}, I_{it}, T_{it})$. We will use centered and normalized embeddings:

$$X_{it}^e = \frac{E_{it} - \frac{1}{n}\sum_i E_{it}}{\|E_{it} - \frac{1}{n}\sum_i E_{it}\|}.$$

Each product is then a point on the sphere.

Does AI Understand Products?

We can evaluate this in different ways:

1. **Qualitative.** Examine clusters of products with similar embeddings, and check if AI-similarity agrees with HI-similarity.
2. **Quantitative.** Check whether AI-embeddings improve predictions of price and quantity signals.

Both tasks are very important for demand analysis, including:

- Computation of hedonic prices (prices as functions of characteristics)
- Hedonic inflation price indices
- Demand and price prediction for new products
- Predicting changes in demand due to changes in price

PCAs and Cluster Visualizations

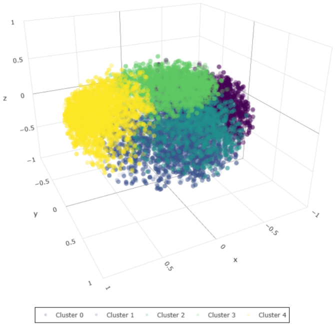


Figure 3: 3D-representation of product embeddings (with image) and five clusters

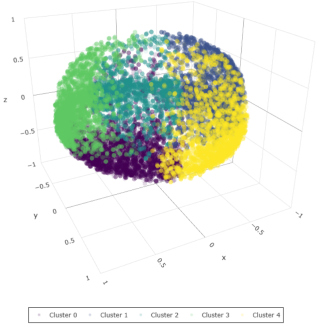


Figure 4: 3D-representation of product embeddings (no image) and five clusters

Cluster Centroids (with Images)

Cluster 0

Cluster 2

Cluster 4

Table 2: Examples of closest products to cluster centers (Tabular + Text + Image); Only three examples of clusters shown.

Cluster Centroids (No Images)

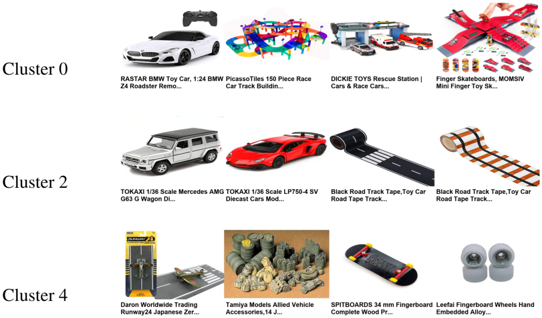


Table 3: Examples of closest products to cluster centers (Tabular + Text); Only three examples of clusters shown.

Gen AI Summary of Cluster Centroids

Cluster 0	Movie-Themed Character Collectibles and Customization Kits	
Cluster 1	Realistic Large-Scale Utility Trucks with Functional Components	
Cluster 2	1:64 Scale Die-Cast Real-World Automotive Replicas	
Cluster 3	Medium-Scale Pull-Back Models and Wooden Train Sets	
Cluster 4	Interactive Stunt Vehicles, Launchers, and Adventure Playsets	

Table 4: Generative AI summaries and images for the five cluster centroids (Tabular + Text + Image).

Quantitative Assessment 1: Hedonics

Objects of Interest:

$$E[Y \mid X], Y \in \{Q, P, \Delta Q, \Delta P\}.$$

These are reduced-form regressions that we also use to fine-tune the embeddings.

Table 5: Test R^2 scores for predicting quantity and price signals.

Method [features] \Target	Q_{it}	P_{it}	ΔQ_{it}	ΔP_{it}
Linear Reg [all tabular]	20.84%	15.82%	7.08%	0.64%
Boosted Trees Reg [all tabular]	47.51%	18.54%	15.13%	0.00%
Deep Learning Reg [text only; invariant tabular]	50.31%	65.74%	9.48%	1.24%
Deep Learning Reg [image and text; invariant tabular]	50.44%	67.26%	10.89%	1.15%
Deep Learning Reg [text only; all tabular]	60.65%	64.89%	14.48%	1.01%
Deep Learning Reg [image and text; all tabular]	59.53%	66.69%	12.17%	1.65%

Quantitative Assessment 2: Dimension Reduction

- **Principal Components (PCA):**

$$X_{i,k}^{pc} := \gamma_k^T X_i^e, \qquad X_i^{pc} := \left(X_{i,k}^{pc}\right)_{k=1}^{K},$$

where γ_k is the k-th eigenvector of the covariance matrix of X_i^e, corresponding to the k-th largest eigenvalue.

- **Centroid Similarities (CS):**

$$X_{i,k}^{sim} := c_k^T X_i^e, \qquad X_i^{sim} := \left(X_{i,k}^{sim}\right)_{k=1}^{K},$$

where c_k is the centroid (mean) of the k-th cluster identified by k-means.

We believe that second approach is more interpretable in our case. What do you think?

5 Dimension-Reduced Features is Good Enough

- 5 centroid similarities or 5 PCAs seem to work nearly as well as 256-dimensional embeddings.
- Enables simpler downstream causal analysis, parsimonious models of elasticity.

Table 6: Test R^2 Scores for ML methods using DL-based PCAs and similarities together with tabular controls.

Method [+ DL Features] \Target	Q_{it}	P_{it}	ΔQ_{it}	ΔP_{it}
Linear Reg [+5 PCAs]	46.47%	60.63%	7.06%	0.64%
Linear Reg [+ 5 Similarities]	47.27%	57.02%	7.07%	0.64%
Linear Reg [+256 Embeddings]	51.24%	62.14%	5.47%	0.00%
Boosted Trees Reg [+5 PCAs]	53.22%	64.36%	13.79%	0.00%
Boosted Trees Reg [+5 Similarities]	52.58%	62.53%	14.63%	0.00%
Boosted Trees Reg [+256 Embeddings]	55.28%	65.64%	10.11%	0.00%

Summarize findings so far:

- AI-generated representations, especially fine-tuned ones, "understand" products.
- Representations provide remarkable boosts in the predictive performance of price and quantity signals.

Causal Inference Next: Price Sensitivity

Goal: **Causal effect of changing the log price on log inverse ranking**

- Begin with the predictive model:

$$L[Q_{it}|P_{it}, X_{it}] = \delta P_{it} + g_t(X_{it})$$

- $L[Y|P, X]$ denotes the projection of Y onto the space of partially linear prediction rules $aP + g(X)$.

Price Sensitivity: Attempt 1

Goal: **Causal effect of changing the log price on log inverse ranking**

- Predictive model: $L[Q_{it}|P_{it}, X_{it}] = \delta P_{it} + g_t(X_{it})$
- The estimated price sensitivity ("elasticity") is very small:

$$\delta \approx [0, -0.2]$$

regardless of the strategies for X or g. This does not make sense from a causal point of view. Why?

Consider Instead Structural Model:

$$Q_{it} = \delta P_{it} + g_t(S_{it}) + \epsilon_{it},$$

where $S_{i,t}$ includes the (1) past sales/visibility/quality of the product and (2) past prices. Omitting these from the model creates a huge omitted variable bias.

A Causal DAG

$$
\begin{tikzcd}
\cdots & S_{i,t-1} \arrow[r] \arrow[rr, bend left] \arrow[dr] & Q_{i,t-1} \arrow[r] & S_{it} \arrow[r] \arrow[rr, bend left] \arrow[dr] & Q_{it} \arrow[r] & S_{i,t+1} & \cdots \\
 & & P_{i,t-1} \arrow[u] \arrow[ur] & & P_{it} \arrow[u] \arrow[ur] & &
\end{tikzcd}
$$

Figure 5: A directed acyclic graph for the dynamic model.

Discussion of limitations later!

Price Sensitivity: Attempt 2

Goal: **Causal effect of changing the log price on log inverse ranking**

- Consider the predictive model using autoregressive price and quantity signals:

$$L[Q_{it}|P_{it}, S_{it}] = \delta P_{it} + g_t(S_{it}),$$

  with $S_{it} = (Q_{i,t-1}, P_{i,t-1}, X_{it})$.
- We obtain "elasticity" estimates:

$$\delta \approx -0.7$$

- To gauge demand elasticity (instead of inverse rank elasticity), multiply by 2. This seems quite plausible.

Price Sensitivity: Attempt 2

Table 7: Estimated price effects based on the partially linear dynamic model.

Specification of Control Function (State S_t)	coef	std err	t	P-val.	[5.0%,	95.0%]
I-1. Linear (P_{t-1}, Q_{t-1})	-0.690	0.040	-17.248	<0.001	-0.756	-0.624
I-1. Linear (P_{t-1}, Q_{t-1}, X^e,X_t^o)	-0.712	0.039	-18.364	<0.001	-0.776	-0.649
I-1. Linear (P_{t-1}, Q_{t-1}, X^{sim},X_t^o)	-0.723	0.039	-18.725	<0.001	-0.786	-0.659
I-2. Linear with Interactions (P_{t-1}, Q_{t-1}, X^{sim},X_t^o)	-0.727	0.039	-18.698	<0.001	-0.791	-0.662
I-3. Boosted Trees (P_{t-1}, Q_{t-1}, X^e,X_t^o)	-0.697	0.049	-14.362	<0.001	-0.704	-0.542
I-3. Boosted Trees (P_{t-1}, Q_{t-1}, X^{sim},X_t^o)	-0.691	0.041	-17.051	<0.001	-0.777	-0.617

Note: Standard errors are clustered at the product level. The LR models are estimated using OLS. The PLR model is estimated using DML with cross-fitted boosted trees.

Price Sensitivity: Attempt 3

The heterogeneous effect model:

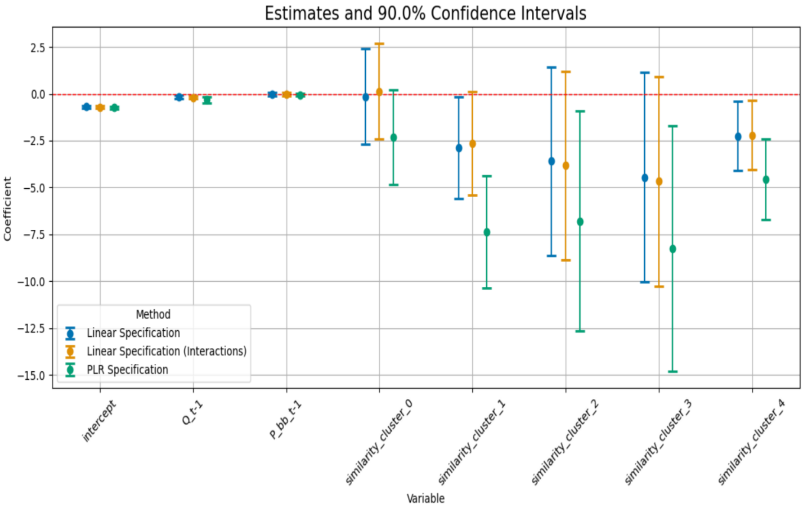
$$L[Q_{it} \mid P_{it}, S_{it}] = \delta_t(S_{i,t})P_{it} + g_t(S_{it}).$$

We make the elasticity a linear function of cluster centroids, price and quantities in the base period:

$$\delta_t(S_{i,t}) = a_0 + \sum_{c=1}^{5} a_c \operatorname{sim}(X_{it}, x_c) + b_1 P_{i,t-1} + b_2 Q_{i,t-1},$$

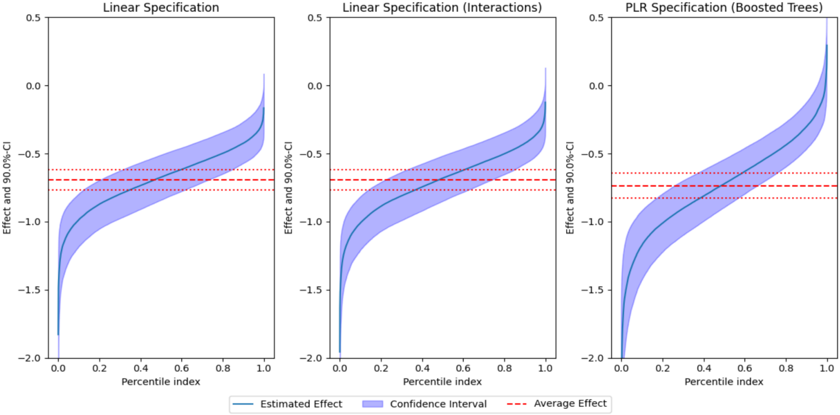
The idea is that price sensitivity might be higher for more visible and more expensive items.

Price Sensitivity: Attempt 3



Price Sensitivity: Attempt 3

The price sensitivity function δ_t in one plot (sorted effects):



Price Sensitivity: Attempt 3

Table 9: p-values for χ^2-test of joint significance of price effect modifiers

Model	All Modifiers	Similarities Only
II.1 Linear Specification	<0.001	0.030
II.2 Linear Specification (Interactions)	<0.001	0.054
II.3 PLR Specification (Boosted Trees)	<0.001	0.001

Note: Standard errors are clustered at the product level. LR is estimated using OLS. The PLR model is estimated using DML with cross-fitted boosted trees (sample size: $38,041$).

Price Sensitivity: Limitations

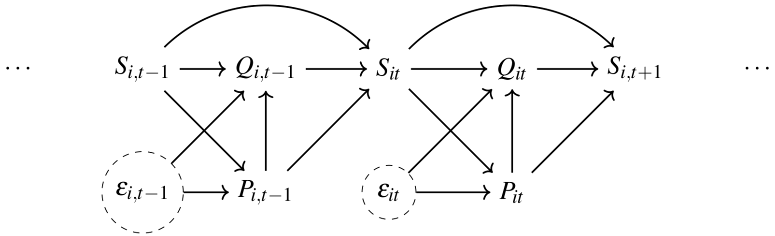


Figure 8: Dynamic model with demand shocks as omitted confounders.

Price Sensitivity: Limitations

- **Endogeneity concern:** Price may respond to demand shocks (ε_{it}), making ε_{it} a confounder. In practice, prices tend to be "sticky," so this channel may be weak.
- **Identification condition:** If the link $\varepsilon_{it} \to P_{it}$ is negligible, then causal interpretation holds. Otherwise, estimates are treated as approximations – see Chernozhukov, Cinelli, Newey et al., 2021 for sharp nonparametric OMVB and bounds.
- **Instrumental variables:** A valid instrument Z_{it} can induce exogenous variation in P_{it}, allowing identification of causal effects of prices on quantities. Follows the classical approach of Wright (1928, republished 2024).
- **Limitations:** No strong instruments found in this setting. Using lagged $Q_{i,t-1}$ is possible, but risks over-interpreting autoregressive structure.

Sensitivity Analysis: Omitted Confounders

- **Concern:** latent demand shocks may affect both price and quantity (endogeneity).
- We use sensitivity bounds for omitted variables (Chernozhukov, Cinelli, Newey et al., 2021), parameterized by the fraction of *residual* variation an unobserved confounder can explain in both Q_{it} and P_{it}.
- If a confounder explains up to 5% of residual variation in both variables, the average rank-elasticity is bounded as

  $$\delta \in [-0.919,\ -0.475],$$

  with one-sided 90% bounds about $[-0.979,\ -0.409]$.
- Bounds include 0 only if the confounder explains roughly 15% of residual variation in both variables.

Conclusion

- AI-generated representations, especially fine-tuned ones, really existing products.
- Representations provide remarkable boosts in prediction of price and quantity signals.
- These representations play key role in predicting price sensitivities/elasticities.
- Probably useful for other causal inference tasks.