Causal Inference for Asset Pricing

Valentin Haddad Zhiguo He Paul Huebner Peter Kondor Erik Loualiche

UCLA, Stanford, SSE, LSE, Minnesota

July 2025

Causal Inference for Asset Pricing

This paper: A framework for causal inference with asset prices and quantities

Causal Inference for Asset Pricing

This paper: A framework for causal inference with asset prices and quantities

- Natural experiments: treatment & control, IV, ...
  - index inclusions, Fed asset purchases, mutual fund reclassifications, ...
  - future of our field: Coppola, Dos Santos, Lu, Mainardi, Selgrad, Siani, Wiegand, ...
- Quantitative demand systems
  - Koijen Yogo 2019, Haddad Huebner Loualiche 2025, ...

Causal Inference for Asset Pricing

This paper: A framework for causal inference with asset prices and quantities

- Natural experiments: treatment & control, IV, ...
  - index inclusions, Fed asset purchases, mutual fund reclassifications, ...
  - future of our field: Coppola, Dos Santos, Lu, Mainardi, Selgrad, Siani, Wiegand, ...

- Quantitative demand systems
  - Koijen Yogo 2019, Haddad Huebner Loualiche 2025, ...

How do those approaches account for substitution and spillovers across assets?

- Traditional methods: "everything is connected," Euler equation tests, factor models, ...

Causal Inference for Asset Pricing

This paper: A framework for causal inference with asset prices and quantities

- Natural experiments: treatment & control, IV, ...
  - index inclusions, Fed asset purchases, mutual fund reclassifications, ...
  - future of our field: Coppola, Dos Santos, Lu, Mainardi, Selgrad, Siani, Wiegand, ...

  How to interpret estimates? Implicit assumptions on spillovers?

- Quantitative demand systems
  - Koijen Yogo 2019, Haddad Huebner Loualiche 2025, ...

  Which results are robust outside of these models and which are specific to these structures?

How do those approaches account for substitution and spillovers across assets?

- Traditional methods: "everything is connected," Euler equation tests, factor models, ...

Question

How do an investor's portfolio decisions respond to prices?

Question

How do an investor's portfolio decisions respond to prices?

$$\underbrace{\mathcal{E}}_{N\times N} = \frac{\partial D}{\partial P} = \left[\frac{\partial D_i}{\partial P_j}\right]_{ij}$$

Elasticity matrix: sensitivity of demand to prices

Question

How do an investor's portfolio decisions respond to prices?

$$\underbrace{\mathcal{E}}_{N\times N} = \frac{\partial D}{\partial P} = \left[\frac{\partial D_i}{\partial P_j}\right]_{ij}$$

Elasticity matrix: sensitivity of demand to prices

- Defined in any theory
- Could be log, levels, shares, changes or not, ...
- Flipside: price impact, how do shifts in demand affect prices?

An Example: CalPERS and Corporate Bonds

- Prices have moved **and no other news**. CalPERS adjusts its bond portfolio:

	Price change	Change in position
1. 10-yr Ford	+ 5%	sell 200
2. 10-yr GM	+ 2%	sell 100
3. 5-yr First Solar	- 1%	buy 100
⋮	⋮	⋮

An Example: CalPERS and Corporate Bonds

- Prices have moved **and no other news**. CalPERS adjusts its bond portfolio:

	Price change	Change in position
1. 10-yr Ford	+ 5%	sell 200
2. 10-yr GM	+ 2%	sell 100
3. 5-yr First Solar	- 1%	buy 100
⋮	⋮	⋮

$$\Delta D_1 = \underbrace{\mathcal{E}_{11}\Delta P_1}_{\text{became more expensive}} + \underbrace{\mathcal{E}_{12}\Delta P_2}_{\text{substitutes from GM}} + \underbrace{\sum_{k\geq 3} \mathcal{E}_{1k}\Delta P_k}_{\text{substitutes from First Solar, ...}}$$

An Example: CalPERS and Corporate Bonds

- Prices have moved **and no other news**. CalPERS adjusts its bond portfolio:

	Price change	Change in position
1. 10-yr Ford	+ 5%	sell 200
2. 10-yr GM	+ 2%	sell 100
3. 5-yr First Solar	- 1%	buy 100
⋮	⋮	⋮

$$\Delta D_1 = \mathcal{E}_{11} \Delta P_1 + \mathcal{E}_{12} \Delta P_2 + \sum_{k \geq 3} \mathcal{E}_{1k} \Delta P_k$$

$$\Delta D_2 = \mathcal{E}_{22} \Delta P_2 + \mathcal{E}_{21} \Delta P_1 + \sum_{k \geq 3} \mathcal{E}_{2k} \Delta P_k$$

$$\vdots \qquad \vdots$$

An Example: CalPERS and Corporate Bonds

- Prices have moved **and no other news**. CalPERS adjusts its bond portfolio:

	Price change	Change in position
1. 10-yr Ford	+ 5%	sell 200
2. 10-yr GM	+ 2%	sell 100
3. 5-yr First Solar	- 1%	buy 100
⋮	⋮	⋮

$$\Delta D_1 = \mathcal{E}_{11} \Delta P_1 + \mathcal{E}_{12} \Delta P_2 + \sum_{k \geq 3} \mathcal{E}_{1k} \Delta P_k$$

$$\Delta D_2 = \mathcal{E}_{22} \Delta P_2 + \mathcal{E}_{21} \Delta P_1 + \sum_{k \geq 3} \mathcal{E}_{2k} \Delta P_k$$

$$\vdots \qquad \vdots$$

$\rightarrow$ **Stuck without additional assumptions**

Making Progress

Two broad paths:

- *Structural approach:* choose a microfoundation and estimate the corresponding model
- *Causal inference:* impose elementary restriction keeping as much flexibility on mechanism as possible while letting the data speak

Making Progress

Two broad paths:

- *Structural approach:* choose a microfoundation and estimate the corresponding model
- *Causal inference:* impose elementary restriction keeping as much flexibility on mechanism as possible while letting the data speak

- **Canonical assumption (SUTVA)**: demand for each asset only depends on its own price $D_i(P_i)$, diagonal $\mathcal{E}$

Making Progress

Two broad paths:

- *Structural approach:* choose a microfoundation and estimate the corresponding model

- *Causal inference:* impose elementary restriction keeping as much flexibility on mechanism as possible while letting the data speak

- ~~**Canonical assumption (SUTVA):** demand for each asset only depends on its own price~~ $\cancel{D_i(P_i)}$~~, diagonal~~ $\cancel{\mathcal{E}}$
  $\rightarrow$ *Portfolio choice, not asset choice!*

Making Progress

Two broad paths:

- *Structural approach:* choose a microfoundation and estimate the corresponding model
- *Causal inference:* impose elementary restriction keeping as much flexibility on mechanism as possible while letting the data speak
- **Our assumption: homogeneous substitution conditional on observables**
  - When the price of a given bond moves, CalPERS changes positions in other bonds based on their observables (e.g. duration, greenness) only

Homogeneous Substitution Conditional on Observables

	Initial position	Price change	New position
1. 10-yr Ford	1,500	+ 5%	$\downarrow$ 1,300
2. 10-yr GM	1,500	+ 2%	$\downarrow$ 1,400
3. 5-yr First Solar	2,000	- 1%	$\uparrow$ 2,100
$\vdots$	$\vdots$	$\vdots$	$\vdots$

$$\Delta D_1 = \mathcal{E}_{11}\Delta P_1 + \mathcal{E}_{12}\Delta P_2 + \sum_{k\geq 3}\mathcal{E}_{1k}\Delta P_k$$

$$\Delta D_2 = \mathcal{E}_{22}\Delta P_2 + \mathcal{E}_{21}\Delta P_1 + \sum_{k\geq 3}\mathcal{E}_{2k}\Delta P_k$$

Homogeneous Substitution Conditional on Observables

	Initial position	Price change	New position
1. 10-yr Ford	1,500	+ 5%	$\downarrow$ 1,300
2. 10-yr GM	1,500	+ 2%	$\downarrow$ 1,400
3. 5-yr First Solar	2,000	- 1%	$\uparrow$ 2,100
$\vdots$	$\vdots$	$\vdots$	$\vdots$

$$\Delta D_1 = \mathcal{E}_{11}\Delta P_1 + \mathcal{E}_{12}\Delta P_2 + \sum_{k\geq 3} \mathcal{E}_{1k}\Delta P_k$$

$$\Delta D_2 = \mathcal{E}_{22}\Delta P_2 + \mathcal{E}_{21}\Delta P_1 + \sum_{k\geq 3} \underbrace{\mathcal{E}_{2k}}_{=\mathcal{E}_{1k}}\Delta P_k$$

- **Compare bonds with same observables:** Ford vs. GM
  - E.g.: CalPERS adjusts Ford and GM equally in response to price of First Solar $\mathcal{E}_{13} = \mathcal{E}_{23}$

Homogeneous Substitution Conditional on Observables

	Initial position	Price change	New position
1. 10-yr Ford	1,500	+ 5%	↓ 1,300
2. 10-yr GM	1,500	+ 2%	↓ 1,400
3. 5-yr First Solar	2,000	- 1%	↑ 2,100
⋮	⋮	⋮	⋮

$$\Delta D_1 = \mathcal{E}_{11}\Delta P_1 + \mathcal{E}_{12}\Delta P_2 + \sum_{k\geq 3} \mathcal{E}_{1k}\Delta P_k$$

$$\Delta D_2 = \mathcal{E}_{22}\Delta P_2 + \mathcal{E}_{21}\Delta P_1 + \sum_{k\geq 3} \underbrace{\mathcal{E}_{2k}}_{=\mathcal{E}_{1k}}\Delta P_k$$

- **Compare bonds with same observables:** Ford vs. GM
  - E.g.: CalPERS adjusts Ford and GM equally in response to price of First Solar $\mathcal{E}_{13} = \mathcal{E}_{23}$

$\rightarrow$ *comparing assets with same observables differences out substitution*

Formal Setup

- **Homogeneous substitution conditional on observables X**

$$\boxed{\mathcal{E}_{il} = \mathcal{E}_{jl} \text{ if } X_i = X_j} \quad \text{for all } i, j \in \mathcal{S}, \text{ and } l \neq i, j,$$

  - If price of 3rd asset move, response of demand for 2 assets with same observables is the same
  - Parametrize linearly: $\mathcal{E}_{il} = \mathcal{E}_{\text{cross}}(X_i, X_l) = X_i' \boldsymbol{\mathcal{E}}_X X_l$

Formal Setup

- **Homogeneous substitution conditional on observables X**

$$\boxed{\mathcal{E}_{il} = \mathcal{E}_{jl} \text{ if } X_i = X_j} \quad \text{for all } i, j \in \mathcal{S}, \text{ and } l \neq i, j,$$

  - If price of 3rd asset move, response of demand for 2 assets with same observables is the same
  - Parametrize linearly: $\mathcal{E}_{il} = \mathcal{E}_{\text{cross}}(X_i, X_l) = X_i' \boldsymbol{\mathcal{E}}_X X_l$

- **Decomposition of demand elasticity:**

$$\begin{aligned} \boldsymbol{\mathcal{E}} &= \text{relative elasticity} + \text{substitution} \\ &= \text{diagonal matrix} + X \underbrace{\boldsymbol{\mathcal{E}}_X}_{K \times K} X' \end{aligned}$$

Formal Setup

- **Homogeneous substitution conditional on observables X**

$$\boxed{\mathcal{E}_{il} = \mathcal{E}_{jl} \text{ if } X_i = X_j} \quad \text{for all } i, j \in \mathcal{S}, \text{ and } l \neq i, j,$$

  - If price of 3rd asset move, response of demand for 2 assets with same observables is the same
  - Parametrize linearly: $\mathcal{E}_{il} = \mathcal{E}_{\text{cross}}(X_i, X_l) = X_i' \boldsymbol{\mathcal{E}}_X X_l$

- **Decomposition of demand elasticity:**

$$\boldsymbol{\mathcal{E}} = \text{relative elasticity} + \text{substitution}$$
$$= \underbrace{\hat{\mathcal{E}}}_{\text{scalar}} I + X \underbrace{\boldsymbol{\mathcal{E}}_X}_{K \times K} X'$$

  - Assume constant relative elasticity $\hat{\mathcal{E}}$ for simplicity, relax in the paper

Many Potential Mechanisms

Key question: What do investors consider when substituting between assets?

- *Investor manages portfolio-level statistic, so substitution depends on asset i's contribution*

Many Potential Mechanisms

Key question: What do investors consider when substituting between assets?

- *Investor manages portfolio-level statistic, so substitution depends on asset i's contribution*

This matters for what observables X_i to include

- *Broad categories*: X_i are group dummies say on durations or industries

Many Potential Mechanisms

Key question: What do investors consider when substituting between assets?

- *Investor manages portfolio-level statistic, so substitution depends on asset i's contribution*

This matters for what observables X_i to include

- *Broad categories*: X_i are group dummies say on durations or industries
- *Risk based motives*: care about portfolio-level factor exposure, so X_i are factor loadings or characteristics that proxy for them
  - Markowitz: $D = \frac{1}{\gamma}\Sigma^{-1}(\mu - P) \Rightarrow \boldsymbol{\mathcal{E}} = -\frac{1}{\gamma}\Sigma^{-1}$
  - If Σ has factor structure: idio risk drives $\hat{\mathcal{E}}$, factor risk drives $\boldsymbol{\mathcal{E}}_X$

Many Potential Mechanisms

Key question: What do investors consider when substituting between assets?

- *Investor manages portfolio-level statistic, so substitution depends on asset i's contribution*

This matters for what observables X_i to include

- *Broad categories*: X_i are group dummies say on durations or industries
- *Risk based motives*: care about portfolio-level factor exposure, so X_i are factor loadings or characteristics that proxy for them
- *Non-risk motives*: X_i is asset weight in this objective

$$\max_D \quad D'(\mu - P) - \frac{\gamma}{2} D'\Sigma D - \frac{\kappa}{2}\left(D'X^{(1)}\right)^2$$

$$\text{such that} \quad D'X^{(2)} \leq \Theta$$

  - Binding constraints (leverage), regulatory score (capital ratio), or stakeholders pressure (greenness)

Cross-Sectional Identification

- **Data-Generating-Process:** Elasticity matrix $\mathcal{E}$ + *homogeneous substitution conditional on observable* X

$$\Delta \mathbf{D} = \mathcal{E} \Delta \mathbf{P} + \epsilon$$

  - Demand shift ϵ correlated with prices: Ford is more expensive because the new F150 is amazing, change in CalPERS financial health, ...

Cross-Sectional Identification

- **Data-Generating-Process:** Elasticity matrix $\boldsymbol{\mathcal{E}}$ + *homogeneous substitution conditional on observable* X

$$\Delta\mathbf{D} = \boldsymbol{\mathcal{E}}\Delta\mathbf{P} + \epsilon$$

  – Demand shift ϵ correlated with prices: Ford is more expensive because the new F150 is amazing, change in CalPERS financial health, ...

- $\boxed{\textbf{Proposition 1}}$ Under our assumption, and the *usual exclusion and relevance restrictions*, the IV estimator identifies the **relative elasticity** $\widehat{\mathcal{E}} = \boldsymbol{\mathcal{E}}_{ii} - \boldsymbol{\mathcal{E}}_{ji}$ for $X_i = X_j$

$$\Delta D_i = \widehat{\mathcal{E}}\Delta P_i + \theta' X_i + e_i$$
$$\Delta P_i = \lambda Z_i + \eta' X_i + u_i$$

  with Z_i instrument for prices ($Z_i \perp \epsilon_i | X_i$)

  – E.g.: Fed buys some bonds but not others

Absorbing Substitution

- Key step: coefficient on observables θ absorbs substitution from other assets

$$
\begin{aligned}
\Delta D_i &= \mathcal{E}_{ii} \Delta P_i + \sum_{j \neq i} X_i' \mathcal{E}_X X_j \Delta P_j + \epsilon_i \\
&= \left(\mathcal{E}_{ii} - X_i' \mathcal{E}_X X_i\right) \Delta P_i + \sum_j X_i' \mathcal{E}_X X_j \Delta P_j + \epsilon_i \\
&= \underbrace{\left(\mathcal{E}_{ii} - X_i' \mathcal{E}_X X_i\right)}_{\text{relative elasticity}} \Delta P_i + X_i' \underbrace{\sum_j \mathcal{E}_X X_j \Delta P_j}_{\text{constant across assets, absorbed in } \theta} + \epsilon_i
\end{aligned}
$$

- **Relative elasticity:** difference between own-price and cross-price elasticity for assets with same observables
  - How does the relative demand for Ford and GM respond to their relative price?
  - In large cross-sections with substantial idiosyncratic risk $\approx$ own-price elasticity

Substitution and Its Estimation

Estimating substitution $\mathcal{E}_X$ crucial for many questions:

- How does CalPERS adjust its portfolio when the price of all bonds drops?
- Will CalPERS maintain its green tilt if green bonds become very expensive relative to brown bonds?

Substitution and Its Estimation

Estimating substitution $\mathcal{E}_X$ crucial for many questions:

- How does CalPERS adjust its portfolio when the price of all bonds drops?
- Will CalPERS maintain its green tilt if green bonds become very expensive relative to brown bonds?

$\boxed{\textbf{Proposition 2}}$ **Impossible to identify substitution with the cross-section alone**

$$\Delta D_i = \widehat{\mathcal{E}} \Delta P_i + X_i' \underbrace{\sum_j \mathcal{E}_X X_j \Delta P_j + \epsilon_i}_{\textbf{BOTH absorbed in } \theta}$$

- Coefficient on X_i measures both substitution and shift in demand for observable
  - Does CalPERS reduce its green tilt because of expensive green bonds or weaker environmental priorities?

Estimating Substitution with the Time Series

- **Classic strategy:** construct portfolios sorted on observables, and measure their price and demand (= portfolio tilt)

Estimating Substitution with the Time Series

- **Classic strategy:** construct portfolios sorted on observables, and measure their price and demand (= portfolio tilt)

$$\Delta P_{agg} = \frac{1}{N}\sum_i \Delta P_i, \qquad \Delta D_{agg} = \frac{1}{N}\sum_i \Delta D_i$$

$$\Delta P_X = \frac{1}{N}\sum_i X_i \Delta P_i \qquad \Delta D_X = \frac{1}{N}\sum_i X_i \Delta D_i$$

Estimating Substitution with the Time Series

- **Classic strategy:** construct portfolios sorted on observables, and measure their price and demand (= portfolio tilt)

$$\Delta P_{agg} = \frac{1}{N}\sum_i \Delta P_i, \qquad \Delta D_{agg} = \frac{1}{N}\sum_i \Delta D_i$$

$$\Delta P_X = \frac{1}{N}\sum_i X_i \Delta P_i \qquad \Delta D_X = \frac{1}{N}\sum_i X_i \Delta D_i$$

- **Proposition 3** Regressing portfolio tilt on portfolio price with **time series instruments** identifies substitution $\mathcal{E}_X$

$$\Delta D_{agg,t} = \bar{\mathcal{E}}_{agg} \Delta P_{agg,t} + \bar{\mathcal{E}}_X \Delta P_{X,t} + \epsilon_{agg,t}$$
$$\Delta D_{X,t} = \tilde{\mathcal{E}}_{agg} \Delta P_{agg,t} + \tilde{\mathcal{E}}_X \Delta P_{X,t} + \epsilon_{X,t}$$

  - Effectively only K assets = portfolios
  - E.g. Fed does more or less QE and operation twist over time

Summary

Homogeneous substitution conditional on observables X:

$$\begin{aligned}\mathcal{E} &= \text{relative elasticity} + \text{substitution} \\ &= \hat{\mathcal{E}} I + X \boldsymbol{\mathcal{E}}_X X'\end{aligned}$$

Consistent with many motives: risk, constraints, non-pecuniary preferences, irrational, ...

Identification:

- Relative elasticity: compare similar assets = cross-sectional IV controlling for X
- Substitution: demand for portfolios based on X = time-series portfolio level instruments

What About Logit?

- Koijen Yogo (2019), Koijen Richmond Yogo (2024):
  - $\exists$ factor models where volatility and expected returns depend non-linearly of prices which yield asset demand in the logit form
  - Logit has non-zero substitution and can be estimated from the cross-section alone

What About Logit?

- Koijen Yogo (2019), Koijen Richmond Yogo (2024):
  - $\exists$ factor models where volatility and expected returns depend non-linearly of prices which yield asset demand in the logit form
  - Logit has non-zero substitution and can be estimated from the cross-section alone

- Logit satisfies our assumption, and its parameter can be robustly interpreted as relative elasticity

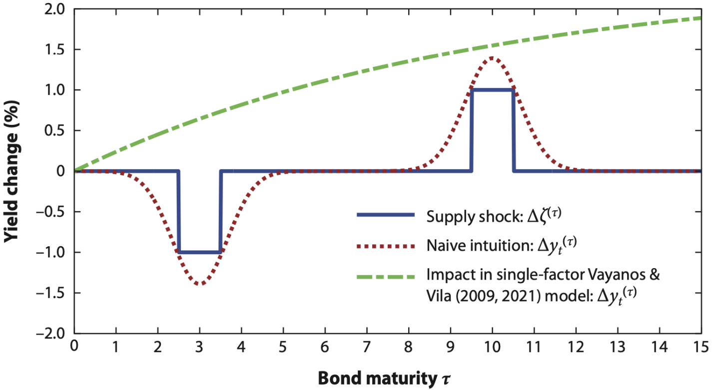
What About Logit?

- Koijen Yogo (2019), Koijen Richmond Yogo (2024):
  - $\exists$ factor models where volatility and expected returns depend non-linearly of prices which yield asset demand in the logit form
  - Logit has non-zero substitution and can be estimated from the cross-section alone

- Logit satisfies our assumption, and its parameter can be robustly interpreted as relative elasticity
- Logit strongly restricts substitution: **an arbitrary factor model is not equivalent to logit**
  - Logit: when the price of any bond $\uparrow$, CalPERS replaces it proportionally to its existing portfolio
  - Factor model: CalPERS replaces it disproportionately with bonds loading on similar factors

Group-based substitution vs Factor Models

- Nested logit (Fang 2023, Koijen Yogo 2024): symmetric groups based on values of observables $\rightarrow$ can use the cross-section of groups to estimate substitution
  - Predict strong local effect and diffuse effect across all other groups
  - Sharply different from factor model with exposure depending on observable (see Cochrane 2008, Vayanos Vila 2021)



Takeaways

- To draw causal inference about demand elasticity, need:
  - A simple assumption: **homogeneous substitution conditional on observables**
    - CalPERS substitutes based on duration and greenness
  - (Standard) source of exogenous variation
    - Fed randomly buys more of some bonds than others, Fed surprisingly engages in QE

- Relative elasticity for similar assets: cross-sectional IV
  - Ford vs GM?
- Substitution = demand for portfolios: time-series IV
  - Green vs brown? Aggregate price?

- Standard structural models of demand rule out most factor-style substitution

Why Causal Inference in Asset Pricing?

- Causal inference particularly valuable when:
  - existing theories are far from the data
  - it is challenging to understand all sources of variations simultaneously
- First step towards better economic theory

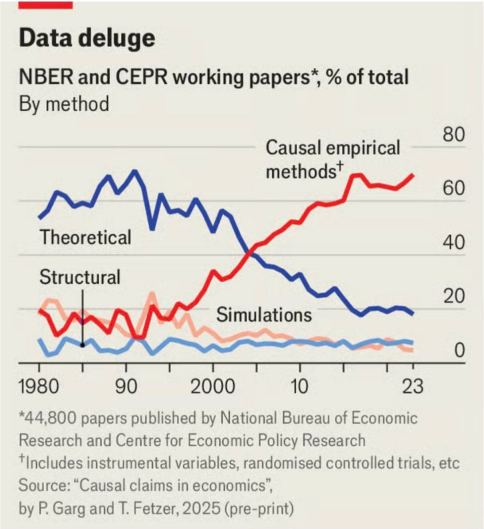


CHART: THE ECONOMIST