Causal Inference for Asset Pricing

Valentin Haddad
UCLA & NBER

Zhiguo He
Stanford & NBER

Paul Huebner
SSE

Péter Kondor
LSE & CEPR

Erik Loualiche
Minnesota

September 2025

Causal Inference for Asset Pricing

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

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, ...
  - many recent JMPs: 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, ...
  - many recent JMPs: 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, ...
  - many recent JMPs: 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, ...

Why is it important?

- Learning from asset quantity data
- Learning from natural experiments
  - Counterpart to growth of micro-empirical methods in macro (e.g. Nakamura Steinsson 2018, Sraer Thesmar 2022)
- Classic models far from the data
- Many important questions are about quantities:
  - Quantitative easing policies (e.g. Haddad Moreira Muir 2025)
  - International capital flows, China and US Treasuries (e.g. Jansen Li Schmid 2025)
  - Rise of passive investing (Haddad, Huebner, Loualiche)

Our Framework

- **Simple portable assumption**: *homogeneous substitution conditional on observables*
  - Diagnostics, empirical design, ...

- **Flexible but parsimonious**: captures the forces of many demand structures, particularly specific to finance
  - Key missing piece of existing models: elasticity to price of factors/characteristics = substitution depends on characteristics

- **Easy estimation**: set of IV/diff-in-diff regression
  - "Separable:" map from different types of natural experiment to different counterfactual
  - Precisely define what is a valid instrument
  - Lots of work on finding instruments, not the focus here

The Problem

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
  - mean-variance: $D = \frac{1}{\gamma}\Sigma^{-1}(M - P)$, $\boldsymbol{\mathcal{E}} = -\frac{1}{\gamma}\Sigma^{-1}$
- Could be log, levels, shares, changes or not, ...
- Flipside: price impact $\mathcal{E}^{-1}$, how do shifts in demand affect prices?

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
  - mean-variance: $D = \frac{1}{\gamma}\Sigma^{-1}(M-P)$, $\mathcal{E} = -\frac{1}{\gamma}\Sigma^{-1}$
- Could be log, levels, shares, changes or not, ...
- Flipside: price impact $\mathcal{E}^{-1}$, how do shifts in demand affect prices?

- *How does CalPERS adjusts its position in 10-year corporate bonds of Ford and GM when their spread changes?*
- *How does AQR move across value and momentum based on their risk premia?*

$\Rightarrow$ **Answer to such questions about different parts of $\mathcal{E}$**

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
⋮	⋮	⋮

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

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

$$\vdots \qquad\qquad \vdots$$

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**

Two Paths

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

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

Two Paths

- **Causal inference:** impose elementary restriction keeping as much flexibility on mechanism as possible while letting the data speak
  - Canonical assumption (SUTVA): when I give you medication, it affects your health but not the control group's health

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

Two Paths

- **Causal inference:** impose elementary restriction keeping as much flexibility on mechanism as possible while letting the data speak
  - Canonical assumption (SUTVA): when I give you medication, it affects your health but not the control group's health
  - Does not work here: demand for each asset only depends on its own price $D_i(P_i) \Rightarrow$ diagonal $\boldsymbol{\mathcal{E}}$

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

Learning from Standard Finance Models

- Assume returns follow a factor structure: exposures β_i and idiosyncratic risk $\sigma^2_{\epsilon,i}$
- *Force 1:* factor management, if expected return only depends on exposures β
  - only buy portfolios replicating the factors (mutual fund theorem)
  - choose exposure to the factors based on the expected returns of those factors
- *Force 2:* "arbitrage": if expected returns deviate from factor pricing
  - buy more of cheap (= high alpha) assets, less of expensive ones

Learning from Standard Finance Models

- Assume returns follow a factor structure: exposures β_i and idiosyncratic risk $\sigma^2_{\epsilon,i}$
- *Force 1:* factor management, if expected return only depends on exposures β
  - only buy portfolios replicating the factors (mutual fund theorem)
  - choose exposure to the factors based on the expected returns of those factors
- *Force 2:* "arbitrage": if expected returns deviate from factor pricing
  - buy more of cheap (= high alpha) assets, less of expensive ones
- A demand formula from Koijen Yogo 2019:

$$\boldsymbol{\mathcal{E}} = -\frac{1}{\gamma}\Sigma^{-1} = \underbrace{-\frac{1}{\gamma\sigma_\epsilon^2}\mathbf{I}}_{\text{diagonal}} + \underbrace{c\boldsymbol{\beta}\boldsymbol{\beta}'}_{\text{substitution matrix}} \qquad \underbrace{c}_{\text{scalar}} = \frac{1}{\gamma\sigma_\epsilon^2}\frac{\sigma_f^2}{\sigma_\epsilon^2 + \sigma_f^2\boldsymbol{\beta}'\boldsymbol{\beta}}$$

Learning from Standard Finance Models

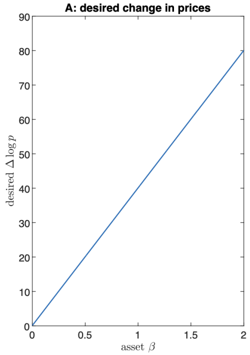
- Assume returns follow a factor structure: exposures β_i and idiosyncratic risk $\sigma^2_{\epsilon,i}$
- *Force 1:* factor management, if expected return only depends on exposures β
  - only buy portfolios replicating the factors (mutual fund theorem)
  - choose exposure to the factors based on the expected returns of those factors
- *Force 2:* "arbitrage": if expected returns deviate from factor pricing
  - buy more of cheap (= high alpha) assets, less of expensive ones
- A demand formula from Koijen Yogo 2019:

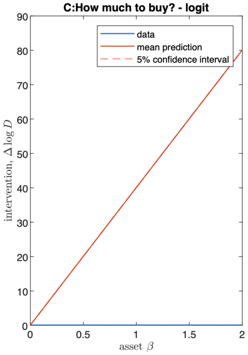
$$\boldsymbol{\mathcal{E}} = -\frac{1}{\gamma}\Sigma^{-1} = \underbrace{-\frac{1}{\gamma\sigma^2_\epsilon}\mathbf{I}}_{\text{diagonal}} + \underbrace{c\boldsymbol{\beta}\boldsymbol{\beta}'}_{\text{substitution matrix}} \qquad \underbrace{c}_{\text{scalar}} = \frac{1}{\gamma\sigma^2_\epsilon}\frac{\sigma^2_f}{\sigma^2_\epsilon + \sigma^2_f\boldsymbol{\beta}'\boldsymbol{\beta}}$$

- **What next?**
  - KY 2019: add additional restrictions, get to logit or similar forms
  - This paper: generalize, keeping only the basic structure of diagonal + substitution driven by observables

Why is everything needed?

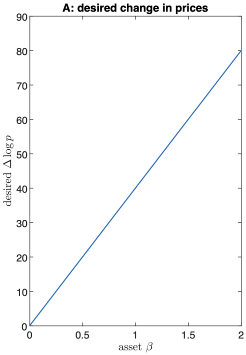
- **Simulation:** Start from factor model demand, increase equally the supply of all assets $\rightarrow$ equilibrium is that price of high beta assets drops more
- **Estimation:** KY 2019 demand and predict what the change in supply was: *erroneously* predict larger increase of supply for high beta assets

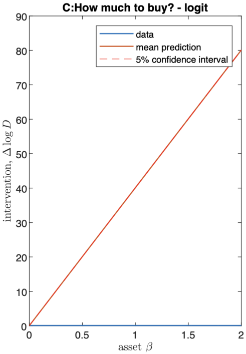


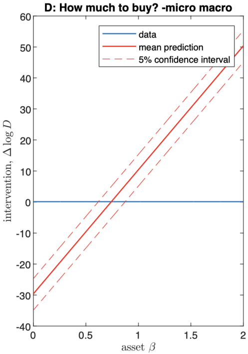


Why is everything needed?

- **Simulation:** Start from factor model demand, increase equally the supply of all assets $\rightarrow$ equilibrium is that price of high beta assets drops more
- **Estimation:** KY 2019 demand and predict what the change in supply was: *erroneously* predict larger increase of supply for high beta assets
  - Adding a macro elasticity (Gabaix Koijen 2025) does not fix it

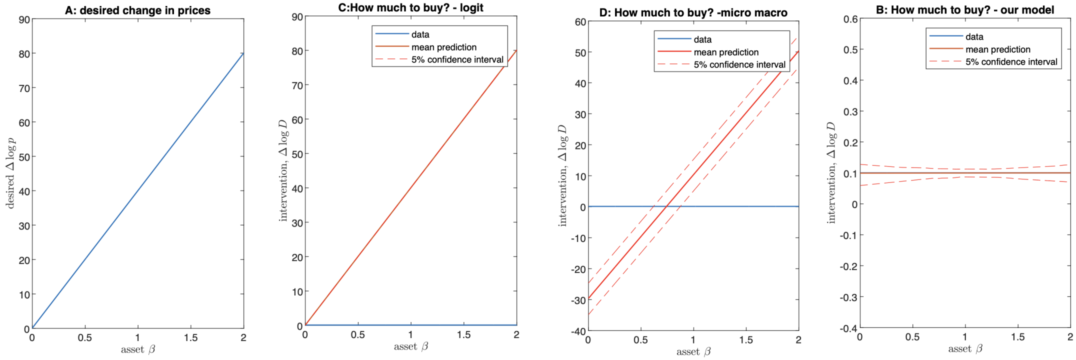






Why is everything needed?

- **Simulation:** Start from factor model demand, increase equally the supply of all assets $\rightarrow$ equilibrium is that price of high beta assets drops more
- **Estimation:** KY 2019 demand and predict what the change in supply was: *erroneously* predict larger increase of supply for high beta assets
  - Adding a macro elasticity (Gabaix Koijen 2025) does not fix it



Framework

Homogeneous Substitution Conditional on Observables

A simple assumption:

- **Homogeneous substitution conditional on observables**
  - CalPERS substitutes across bonds based on their observables (e.g. duration, greenness) only

Homogeneous Substitution Conditional on Observables

A simple assumption:

- **Homogeneous substitution conditional on observables**
  - CalPERS substitutes across bonds based on their observables (e.g. duration, greenness) only

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

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

- **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

A simple assumption:

- **Homogeneous substitution conditional on observables**
  - CalPERS substitutes across bonds based on their observables (e.g. duration, greenness) only

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

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

$$\textit{Diff-in-diff:} \quad \Delta D_1 - \Delta D_2 = \widehat{\mathcal{E}}(\Delta P_1 - \Delta P_2) \text{ if same relative elasticity}$$

- **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:**

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

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

Questions Revisited

$$\mathcal{E} = \text{relative elasticity} + \text{substitution}$$
$$= \widehat{\mathcal{E}} I + X \boldsymbol{\mathcal{E}}_X X'$$

Different questions are about different parts of $\mathcal{E}$

- *How does CalPERS adjusts its position in 10-year corporate bonds of Ford and GM when their spread changes?*
- *How does AQR move across factors based on factor risk premia?*

Questions Revisited

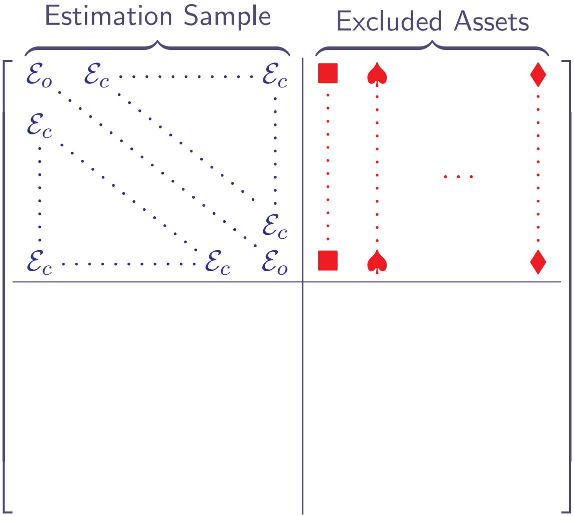
$$\mathcal{E} = \text{relative elasticity} + \text{substitution}$$
$$= \widehat{\mathcal{E}} I + X \boldsymbol{\mathcal{E}}_X X'$$

Different questions are about different parts of $\mathcal{E}$

- *How does CalPERS adjusts its position in 10-year corporate bonds of Ford and GM when their spread changes?*
  - Asset-specific behavior characterized by the relative elasticity $\widehat{\mathcal{E}}$
- *How does AQR move across factors based on factor risk premia?*
  - Question about substitution characterized by $\boldsymbol{\mathcal{E}}_X$

Local Experiments

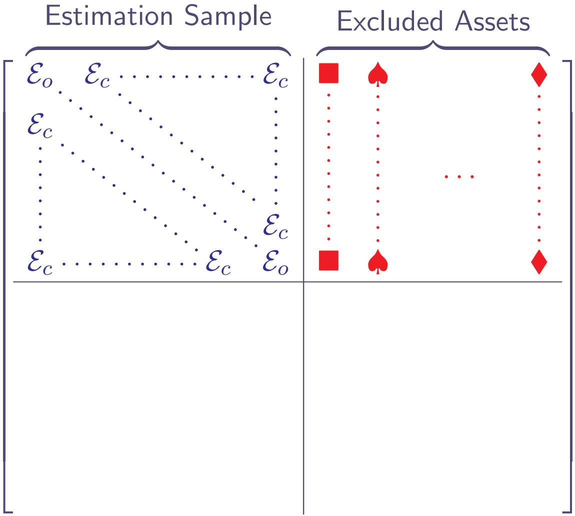
- With few close assets: ignore observables & assume full homogeneity
  - Same own- and cross-price elasticity for every pair of assets in $\mathcal{S}$
  - Demand for all assets in $\mathcal{S}$ responds in same way to price of 5-year First Solar bond (outside $\mathcal{S}$)



Local Experiments

- With few close assets: ignore observables & assume full homogeneity
  - Same own- and cross-price elasticity for every pair of assets in $\mathcal{S}$
  - Demand for all assets in $\mathcal{S}$ responds in same way to price of 5-year First Solar bond (outside $\mathcal{S}$)

- Risk-based models: assets have common variance & covariance + *identical covariance with outside assets*

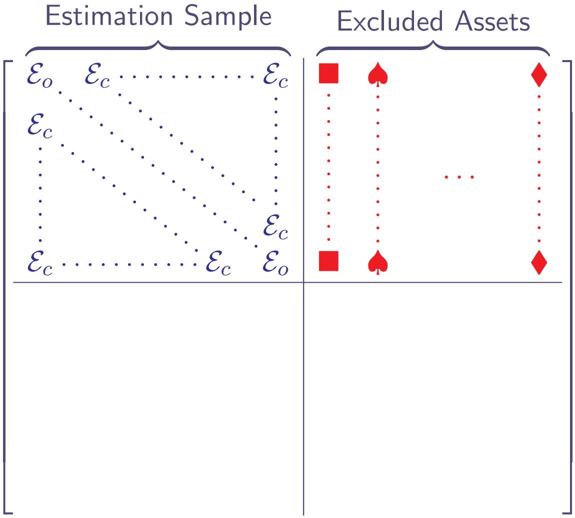


Local Experiments

- With few close assets: ignore observables & assume full homogeneity
  - Same own- and cross-price elasticity for every pair of assets in $\mathcal{S}$
  - Demand for all assets in $\mathcal{S}$ responds in same way to price of 5-year First Solar bond (outside $\mathcal{S}$)

- Risk-based models: assets have common variance & covariance + *identical covariance with outside assets*

- Diagnostic: balance between treated (high Z_i) and control (low Z_i) on covariance with broad factors



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?
  - Useful to answer relative Qs and construct relative counterfactuals
  - In large cross-sections with substantial idiosyncratic risk $\approx$ own-price elasticity
  - What GE theorists call the Morishima elasticity Gabaix Koijen 2025 the micro-elasticity

FAQ

1. *"I exclude some assets from my sample because I don't have data on them. Is this a problem?"* **No**, as long as you assume the **same structure** applies to **excluded assets**. Conditional on X_i, the omitted asset affects others symmetrically, which differences out.

FAQ

1. *"I exclude some assets from my sample because I don't have data on them. Is this a problem?"* **No**, as long as you assume the **same structure** applies to **excluded assets**. Conditional on X_i, the omitted asset affects others symmetrically, which differences out.

2. *"What if I have a repeated cross-section?"* Control for observables X_i **interacted with time fixed effects**.

FAQ

1. *"I exclude some assets from my sample because I don't have data on them. Is this a problem?"* **No**, as long as you assume the **same structure** applies to **excluded assets**. Conditional on X_i, the omitted asset affects others symmetrically, which differences out.

2. *"What if I have a repeated cross-section?"* Control for observables X_i **interacted with time fixed effects**.

3. *"What do I estimate in a price impact regression of exogenous demand shocks on prices?"* Under **our assumption**, the identified coefficient is the **relative multiplier** $-1/\widehat{\mathcal{E}}$.

FAQ

1. *"I exclude some assets from my sample because I don't have data on them. Is this a problem?"* **No**, as long as you assume the **same structure** applies to **excluded assets**. Conditional on X_i, the omitted asset affects others symmetrically, which differences out.

2. *"What if I have a repeated cross-section?"* Control for observables X_i **interacted with time fixed effects**.

3. *"What do I estimate in a price impact regression of exogenous demand shocks on prices?"* Under **our assumption**, the identified coefficient is the **relative multiplier** $-1/\widehat{\mathcal{E}}$.

4. *"How about equilibrium spillovers?"* If the Fed buys a bond but not a similar one, the **price of the similar one also moves**. If LOOP holds, this may threaten instrument relevance (testable), but **not the exclusion restriction**.

FAQ

1. *"I exclude some assets from my sample because I don't have data on them. Is this a problem?"* **No**, as long as you assume the **same structure** applies to **excluded assets**. Conditional on X_i, the omitted asset affects others symmetrically, which differences out.

2. *"What if I have a repeated cross-section?"* Control for observables X_i **interacted with time fixed effects**.

3. *"What do I estimate in a price impact regression of exogenous demand shocks on prices?"* Under **our assumption**, the identified coefficient is the **relative multiplier** $-1/\widehat{\mathcal{E}}$.

4. *"How about equilibrium spillovers?"* If the Fed buys a bond but not a similar one, the **price of the similar one also moves**. If LOOP holds, this may threaten instrument relevance (testable), but **not the exclusion restriction**.

5. *"Can I recover the own-price elasticity from my cross-sectional regression?"* In general, **no** because the **own-price elasticity** combines **both** the relative elasticity and substitution. Cross-sectional regressions only identify part of $\boldsymbol{\mathcal{E}}$.

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**

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

- 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?
  - This is a **missing coefficients** problem

Demand–Price Decomposition

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

Demand–Price Decomposition

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

- X_i (which is normalized) for example captures greenness

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

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

$$\Delta D_{idio,i} = \Delta D_i - \Delta D_{agg} - X_i \Delta D_X \qquad \Delta P_{idio,i} = \Delta P_i - \Delta P_{agg} - X_i \Delta P_X$$

Demand–Price Decomposition

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

- X_i (which is normalized) for example captures greenness

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

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

$$\Delta D_{idio,i} = \Delta D_i - \Delta D_{agg} - X_i \Delta D_X \qquad \Delta P_{idio,i} = \Delta P_i - \Delta P_{agg} - X_i \Delta P_X$$

- *Separate the response of demand to prices into three univariate components:*

Relative: $\Delta D_{idio,i} = \widehat{\mathcal{E}} \Delta P_{idio,i}$

Meso: $\Delta D_X = \tilde{\mathcal{E}}_{agg} \Delta P_{agg} + \tilde{\mathcal{E}}_X \Delta P_X$

Macro: $\Delta D_{agg} = \bar{\mathcal{E}}_{agg} \Delta P_{agg} + \bar{\mathcal{E}}_X \Delta P_X$

Estimating Substitution with the Time Series

- $\boxed{\textbf{Proposition 3}}$ Regressing portfolio tilts on portfolio prices with **time series instruments** identifies substitution $\mathcal{E}_X$

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

  - 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:

$$\mathcal{E} = \text{relative elasticity} + \text{substitution}$$
$$= \widehat{\mathcal{E}} I + X \boldsymbol{\mathcal{E}}_X X'$$

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 inferred 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 inferred from the cross-section alone
- Logit satisfies our assumption, and its parameter can be robustly interpreted as relative elasticity

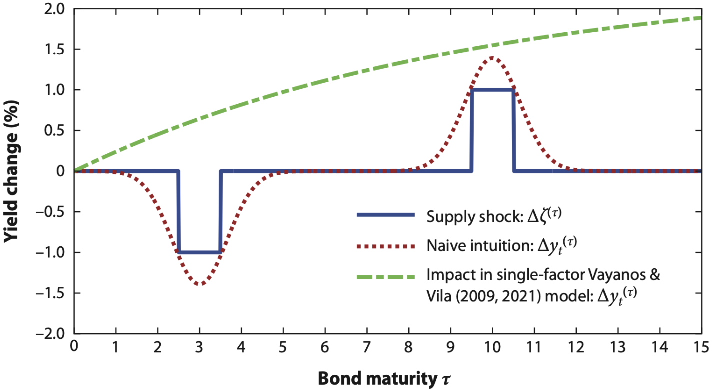
What About Logit?

- 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 inferred 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 → 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)



Example: Corporate Bond Relative Multiplier

- U.S. investment-grade corporate bonds (following Chaudhary Fu Li, 2024)
- Steps to go through when conducting causal inference in asset pricing:
  1. choose a source of variation
  2. assess exogeneity
  3. assess assumptions A1 and A2 and select observables + units
  4. implement the regression analysis
- Step 1: flow-induced demand shock Z_{it}: fund flow in mutual funds $\times$ portfolio composition (Coval Stafford 2007, Lou 2012)
- Step 2: assess exogeneity, i.e., $Z_{it} \perp \epsilon_{it} | X_{it}$
  - example threat to identification: another investor type (e.g., insurance companies) buying the same bonds held by mutual funds that receive a lot of inflows

Step 3: Diagnostic for Homogeneous Substitution – Balance on covariances

Do treated & control bonds comove the same way with broad portfolios?

1. At each date t, form a long-short portfolio based on treatment status
2. Compute the β of the long-short return on broad indices in a window around t (here: 2y)
3. β different from zero $\Rightarrow$ substitution likely not homogeneous

Step 3: Diagnostic for Homogeneous Substitution – Balance on covariances

Do treated & control bonds comove the same way with broad portfolios?

1. At each date t, form a long-short portfolio based on treatment status
2. Compute the β of the long-short return on broad indices in a window around t (here: 2y)
3. β different from zero $\Rightarrow$ substitution likely not homogeneous

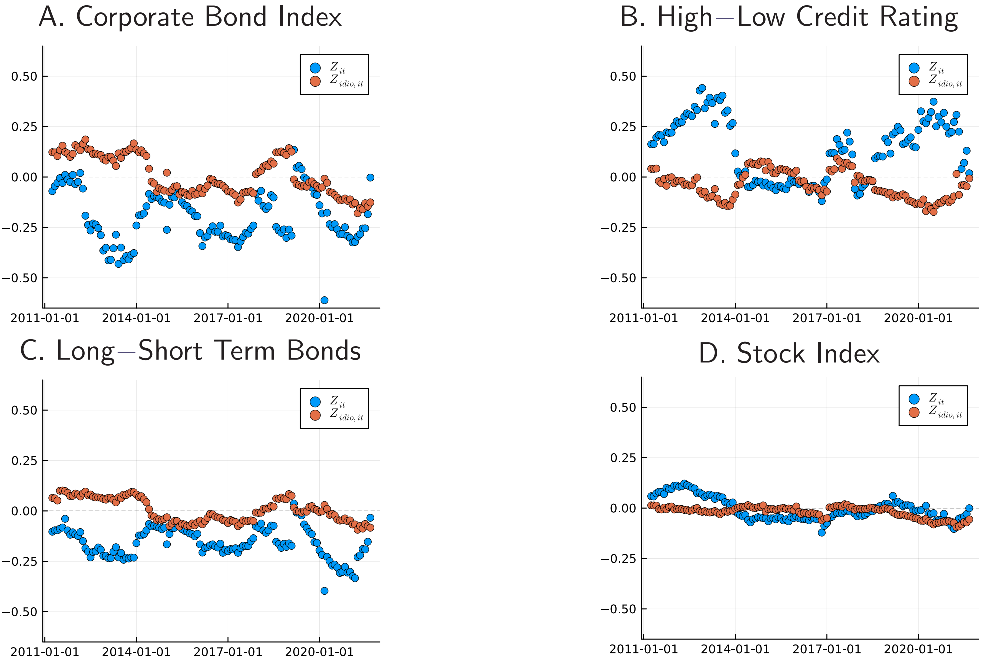
- Treated & control bonds may be differentially exposed to X, driving differences in β $\rightarrow$ natural if investors choose their flows along dimensions like duration and credit risk
- Do they comove the same *conditional on observables*?
- $Z_{idio,it}$: residual of instrument regressed on a date fixed effect, **duration** $\times$ date fixed effects and **credit rating** $\times$ date fixed effects

Step 3: Diagnostic for Homogeneous Substitution – Balance on covariances

Do treated & control bonds comove the same way with broad portfolios?

1. At each date t, form a long-short portfolio based on treatment status
2. Compute the β of the long-short return on broad indices in a window around t (here: 2y)
3. β different from zero $\Rightarrow$ substitution likely not homogeneous

- Treated & control bonds may be differentially exposed to X, driving differences in β $\rightarrow$ natural if investors choose their flows along dimensions like duration and credit risk
- Do they comove the same *conditional on observables*?
- $Z_{idio,it}$: residual of instrument regressed on a date fixed effect, **duration** $\times$ date fixed effects and **credit rating** $\times$ date fixed effects
- Alternative unit to bond returns: yield changes ▸ A1 yield changes ▸ Multiplier yield changes
- Similar diagnostic for constant relative elasticity: balance on idiosyncratic volatility ▸ A2 diagnostic



A. Corporate Bond Index
B. High—Low Credit Rating
C. Long—Short Term Bonds
D. Stock Index
Z_{it}
$Z_{idio,\,it}$
0.50
0.25
0.00
−0.25
−0.50
2011-01-01
2014-01-01
2017-01-01
2020-01-01

Step 4: Implement the regression

Relative multiplier $\widehat{\mathcal{M}} \approx 0$

	Return $\Delta P_{it}/P_{i,t-1}$				
	(1)	(2)	(3)	(4)	(5)
Demand shock:					
Z_{it}	1.541*	-0.254	0.019		
	(0.637)	(0.229)	(0.065)		
$Z_{idio,it}$				0.019	0.019
				(0.065)	(0.065)
Date Fixed Effects		Yes	Yes	Yes	Yes
Duration $\times$ Date Fixed Effects			Yes	Yes	
Credit Rating $\times$ Date Fixed Effects			Yes	Yes	
N	646,335	646,335	646,335	646,335	646,335
R^2	0.010	0.415	0.632	0.632	0.415

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

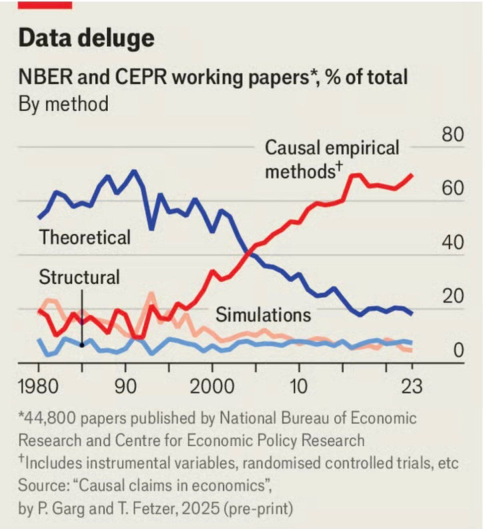
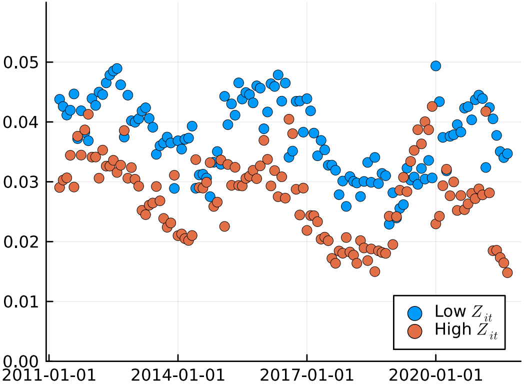


CHART: THE ECONOMIST

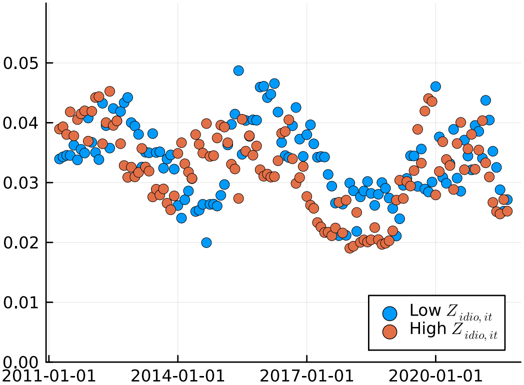
Diagnostic for Constant Relative Elasticity

- **Balance on idiosyncratic volatility**

A. Idiosyncratic Volatility (Z_{it})



B. Idiosyncratic Volatility ($Z_{idio,it}$)

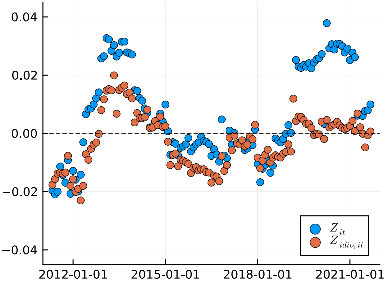


- Average idiosyncratic volatility among treated versus control bonds

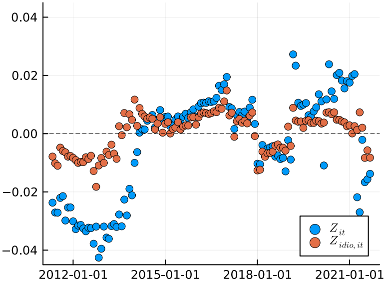
Back

Back

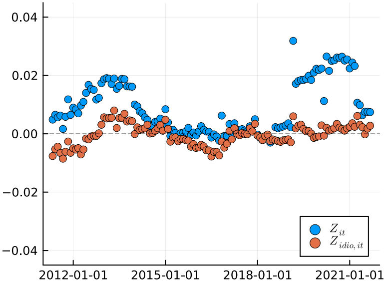
A. Corporate Bond Index



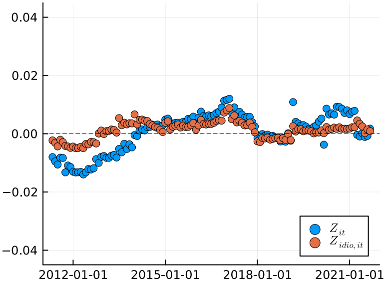
B. High—Low Credit Rating



C. Long—Short Term Bonds



D. Stock Index



▸ Back

Relative multiplier $\widehat{\mathcal{M}} = -0.072$

	Yield change ΔY_{it}				
	(1)	(2)	(3)	(4)	(5)
Demand shock:					
Z_{it}	-0.384*	-0.104*	-0.072**		
	(0.166)	(0.047)	(0.027)		
$Z_{idio,it}$				-0.072**	-0.072**
				(0.027)	(0.027)
Date Fixed Effects		Yes	Yes	Yes	Yes
Duration $\times$ Date Fixed Effects			Yes	Yes	
Credit Rating $\times$ Date Fixed Effects			Yes	Yes	
N	630,255	630,255	630,255	630,255	630,255
R^2	0.004	0.071	0.089	0.089	0.070