# Week 6:
# Instrumental Variables and Event Studies

Dr Christian Engels
*ce50@st-andrews.ac.uk*

FI5699 Dissertation Module

Department of Finance
University of St Andrews Business School

University of
St Andrews

# Introduction

# What We Will Cover Today

This session introduces two powerful tools: instrumental variables (IV) for causal inference under endogeneity, and classical event studies for measuring stock-price reactions to corporate events.

1. **When Fixed Effects Aren't Enough** — time-varying confounders
2. **The IV Idea** — instruments, exclusion, and the Wald estimand
3. **IV Estimation** — 2SLS, weak instruments, first-stage diagnostics
4. **What Does IV Identify?** — LATE and compliers
5. **Running IV in Python** — `pyfixest` three-part formula
6. **Classical Event Studies** — abnormal returns and CARs
7. **Event Studies in Practice** — real data, design choices, and what the literature finds

By the end of today, you will be able to estimate, interpret, and present IV regressions and classical event studies.

# Recap: What You Learned in Week 5

| Concept | What you learned |
|---|---|
| Panel data | Two dimensions: units $\times$ time |
| Fixed effects | Absorb time-invariant unobserved heterogeneity |
| Clustered SEs | Correct inference when residuals correlate within groups |
| `pyfixest` | Formula syntax: `"y ~ x1 + x2 | firm + year"` |

Today we ask two questions: (1) what if the confounders **change over time**? We need instrumental variables. (2) How do stock prices react to specific events? We need event studies.

# When Fixed Effects Aren’t Enough

# Fixed Effects Remove Only Time-Invariant Confounders

Recall the FE model: $Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}$

The unit FE $\alpha_i$ absorbs anything *constant* within a firm. But $\varepsilon_{it}$ still contains time-varying confounders that correlate with $X_{it}$.

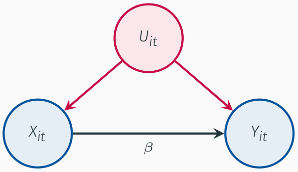
# Fixed Effects Remove Only Time-Invariant Confounders

Recall the FE model: $Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}$

The unit FE $\alpha_i$ absorbs anything *constant* within a firm. But $\varepsilon_{it}$ still contains time-varying confounders that correlate with $X_{it}$.

**Example:**

- $Y$ = firm earnings




The red arrows remain even with FE: time-varying OVB.

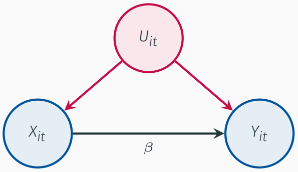
# Fixed Effects Remove Only Time-Invariant Confounders

Recall the FE model: $Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}$

The unit FE $\alpha_i$ absorbs anything *constant* within a firm. But $\varepsilon_{it}$ still contains time-varying confounders that correlate with $X_{it}$.

**Example:**

- $Y$ = firm earnings
- $X$ = military service (Angrist 1990)




The red arrows remain even with FE: time-varying OVB.

# Fixed Effects Remove Only Time-Invariant Confounders

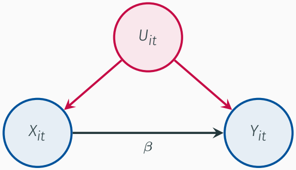
Recall the FE model: $Y_{it} = \alpha_i + \beta X_{it} + \varepsilon_{it}$

The unit FE $\alpha_i$ absorbs anything *constant* within a firm. But $\varepsilon_{it}$ still contains time-varying confounders that correlate with $X_{it}$.

**Example:**

- $Y$ = firm earnings
- $X$ = military service (Angrist 1990)
- Confounder: civilian labour market conditions *at the time of enlistment*

These vary over time — FE cannot absorb them.




The red arrows remain even with FE: time-varying OVB.

# Three Scenarios and Their Solutions

| Problem | Source of bias | Solution |
|---|---|---|
| Time-invariant confounders | $\text{Corr}(\alpha_i, X_{it}) \neq 0$ | Fixed effects |
| Time-varying confounders | $\text{Corr}(U_{it}, X_{it}) \neq 0$ | Instrumental variables |
| Reverse causality | $Y_{it} \rightarrow X_{it}$ | IV or natural experiment |

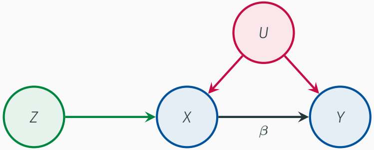
IV gives us a way to isolate **exogenous variation** in $X$ — variation that cannot be contaminated by confounders or reverse causality.

# The IV Idea

# An Instrument Isolates Exogenous Variation in the Treatment

An instrument $Z_i$ is a variable that:

1. **Relevance:** $Z$ is correlated with $X$



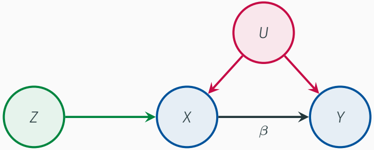

No arrow from $U$ to $Z$ (independence).

No arrow from $Z$ to $Y$ (exclusion).

# An Instrument Isolates Exogenous Variation in the Treatment

An instrument $Z_i$ is a variable that:

1. **Relevance:** $Z$ is correlated with $X$
   $\rightarrow$ $Z$ shifts the treatment



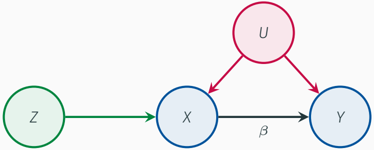

No arrow from $U$ to $Z$ (independence).

No arrow from $Z$ to $Y$ (exclusion).

# An Instrument Isolates Exogenous Variation in the Treatment

An instrument $Z_i$ is a variable that:

1. **Relevance:** $Z$ is correlated with $X$
   $\rightarrow$ $Z$ shifts the treatment
2. **Independence:** $Z$ is as-good-as-randomly assigned



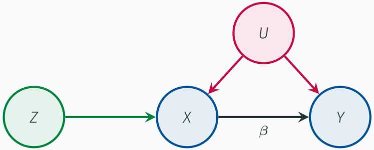

No arrow from $U$ to $Z$ (independence).

No arrow from $Z$ to $Y$ (exclusion).

# An Instrument Isolates Exogenous Variation in the Treatment

An instrument $Z_i$ is a variable that:

1. **Relevance:** $Z$ is correlated with $X$
   $\rightarrow$ $Z$ shifts the treatment
2. **Independence:** $Z$ is as-good-as-randomly assigned
   $\rightarrow$ $Z$ is uncorrelated with confounders



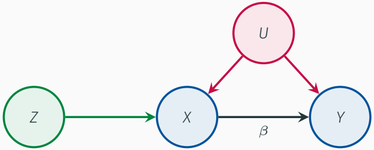

No arrow from $U$ to $Z$ (independence).

No arrow from $Z$ to $Y$ (exclusion).

# An Instrument Isolates Exogenous Variation in the Treatment

An instrument $Z_i$ is a variable that:

1. **Relevance:** $Z$ is correlated with $X$
   $\rightarrow$ $Z$ shifts the treatment
2. **Independence:** $Z$ is as-good-as-randomly assigned
   $\rightarrow$ $Z$ is uncorrelated with confounders
3. **Exclusion:** $Z$ affects $Y$ *only through* $X$



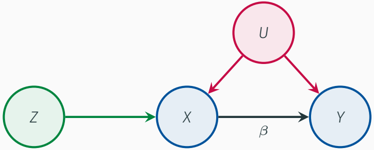

No arrow from $U$ to $Z$ (independence).

No arrow from $Z$ to $Y$ (exclusion).

# An Instrument Isolates Exogenous Variation in the Treatment

An instrument $Z_i$ is a variable that:

1. **Relevance:** $Z$ is correlated with $X$
   $\rightarrow$ $Z$ shifts the treatment
2. **Independence:** $Z$ is as-good-as-randomly assigned
   $\rightarrow$ $Z$ is uncorrelated with confounders
3. **Exclusion:** $Z$ affects $Y$ *only through* $X$
   $\rightarrow$ no direct effect of $Z$ on $Y$




No arrow from $U$ to $Z$ (independence).

No arrow from $Z$ to $Y$ (exclusion).

# Angrist (1990): The Vietnam Draft Lottery as an Instrument

**Question:** What is the effect of military service on lifetime earnings?

| Variable | Definition |
|---|---|
| $Y_i$ | Adult earnings |
| $D_i$ | Military service (0/1) |
| $Z_i$ | Draft eligibility (0/1) |

The draft lottery randomly assigned eligibility:

- Relevance: eligible men more likely to serve

**The Wald estimand:**

$$\beta^{IV} = \frac{E[Y \mid Z = 1] - E[Y \mid Z = 0]}{E[D \mid Z = 1] - E[D \mid Z = 0]}$$

Numerator: reduced form (effect of eligibility on earnings).

Denominator: first stage (effect of eligibility on service probability).

# Angrist (1990): The Vietnam Draft Lottery as an Instrument

**Question:** What is the effect of military service on lifetime earnings?

| Variable | Definition |
|---|---|
| $Y_i$ | Adult earnings |
| $D_i$ | Military service (0/1) |
| $Z_i$ | Draft eligibility (0/1) |

The draft lottery randomly assigned eligibility:

- Relevance: eligible men more likely to serve
- Independence: lottery number is random

**The Wald estimand:**

$$\beta^{IV} = \frac{E[Y \mid Z = 1] - E[Y \mid Z = 0]}{E[D \mid Z = 1] - E[D \mid Z = 0]}$$

Numerator: reduced form (effect of eligibility on earnings).

Denominator: first stage (effect of eligibility on service probability).

# Angrist (1990): The Vietnam Draft Lottery as an Instrument

**Question:** What is the effect of military service on lifetime earnings?

| Variable | Definition |
|---|---|
| $Y_i$ | Adult earnings |
| $D_i$ | Military service (0/1) |
| $Z_i$ | Draft eligibility (0/1) |

The draft lottery randomly assigned eligibility:

- Relevance: eligible men more likely to serve
- Independence: lottery number is random
- Exclusion: lottery only affects earnings through service

**The Wald estimand:**

$$\beta^{IV} = \frac{E[Y \mid Z = 1] - E[Y \mid Z = 0]}{E[D \mid Z = 1] - E[D \mid Z = 0]}$$

Numerator: reduced form (effect of eligibility on earnings).

Denominator: first stage (effect of eligibility on service probability).

# White Veterans Lost $2,000 per Year in Earnings

Angrist (1990) found that white Vietnam veterans earned approximately 15% less than comparable non-veterans, a decade after discharge.

**Two-Sample IV estimates (1981–84 average, 1978$):**

| Earnings measure | IV estimate (SE) |
|---|---|
| FICA taxable | −$1,563 (522) |
| Adjusted FICA | −$1,920 (576) |
| W-2 compensation | −$2,095 (646) |

All estimates exceed twice their standard errors.

**Structural explanation:**

The penalty is consistent with a loss of $\approx$ 2 years of civilian experience (SE 0.38).

At approximately age 50, the veteran penalty vanishes — the experience gap closes.

# IV Reversed the Sign: Selection Bias Was Hiding the True Effect

**OLS** comparisons showed veterans earning *more* than non-veterans.

# IV Reversed the Sign: Selection Bias Was Hiding the True Effect

**OLS** comparisons showed veterans earning *more* than non-veterans.

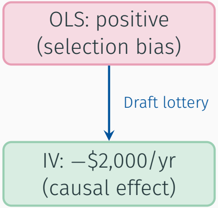
Why? Men with fewer civilian opportunities were more likely to enlist. OLS confounded the treatment effect with this selection.

# IV Reversed the Sign: Selection Bias Was Hiding the True Effect

**OLS** comparisons showed veterans earning *more* than non-veterans.

Why? Men with fewer civilian opportunities were more likely to enlist. OLS confounded the treatment effect with this selection.

**IV** isolated the exogenous variation from the draft lottery — and revealed a large negative effect.




This is why IV matters: it can overturn the conclusion from naive comparisons.

# Every IV Can Be Decomposed into Two Regressions

The IV estimand $\beta^{IV} = \rho/\pi$ is the ratio of two population regressions:

| Regression | Equation | What it shows |
|---|---|---|
| Reduced form | $Y_i = \kappa + \rho Z_i + \nu_i$ | Effect of $Z$ on $Y$ |
| First stage | $D_i = \mu + \pi Z_i + \eta_i$ | Effect of $Z$ on $D$ |

**Always report the reduced form and first stage alongside your IV estimate.**

The reduced form is interpretable on its own as an intent-to-treat effect. The first stage tells you how strongly the instrument shifts the treatment.

# IV Estimation

# 2SLS Uses Only Instrument-Predicted Variation in $X$

Two-stage least squares (2SLS) is the standard IV estimator. With multiple instruments, it combines them efficiently:

1. **First stage:** Regress $X$ on $Z$ and controls $W$. Get fitted values $\hat{X}$.

# 2SLS Uses Only Instrument-Predicted Variation in $X$

Two-stage least squares (2SLS) is the standard IV estimator. With multiple instruments, it combines them efficiently:

1. **First stage:** Regress $X$ on $Z$ and controls $W$. Get fitted values $\hat{X}$.
2. **Second stage:** Regress $Y$ on $\hat{X}$ and controls $W$.

# 2SLS Uses Only Instrument-Predicted Variation in $X$

Two-stage least squares (2SLS) is the standard IV estimator. With multiple instruments, it combines them efficiently:

1. **First stage:** Regress $X$ on $Z$ and controls $W$. Get fitted values $\hat{X}$.
2. **Second stage:** Regress $Y$ on $\hat{X}$ and controls $W$.

# 2SLS Uses Only Instrument-Predicted Variation in $X$

Two-stage least squares (2SLS) is the standard IV estimator. With multiple instruments, it combines them efficiently:

1. **First stage:** Regress $X$ on $Z$ and controls $W$. Get fitted values $\hat{X}$.
2. **Second stage:** Regress $Y$ on $\hat{X}$ and controls $W$.

**Why it works:**

- $\hat{X}$ contains only the variation in $X$ that is *predicted by the instruments*

Never run 2SLS by hand.

Use `pf.feols()` with the IV formula. Hand-computed SEs from two separate regressions are **wrong**.

# 2SLS Uses Only Instrument-Predicted Variation in $X$

Two-stage least squares (2SLS) is the standard IV estimator. With multiple instruments, it combines them efficiently:

1. **First stage:** Regress $X$ on $Z$ and controls $W$. Get fitted values $\hat{X}$.
2. **Second stage:** Regress $Y$ on $\hat{X}$ and controls $W$.

**Why it works:**

- $\hat{X}$ contains only the variation in $X$ that is *predicted by the instruments*
- This predicted variation is exogenous (by the independence of $Z$)

Never run 2SLS by hand.

Use `pf.feols()` with the IV formula. Hand-computed SEs from two separate regressions are **wrong**.

# 2SLS Uses Only Instrument-Predicted Variation in $X$

Two-stage least squares (2SLS) is the standard IV estimator. With multiple instruments, it combines them efficiently:

1. **First stage:** Regress $X$ on $Z$ and controls $W$. Get fitted values $\hat{X}$.
2. **Second stage:** Regress $Y$ on $\hat{X}$ and controls $W$.

**Why it works:**

- $\hat{X}$ contains only the variation in $X$ that is *predicted by the instruments*
- This predicted variation is exogenous (by the independence of $Z$)
- The second stage uses only "clean" variation

Never run 2SLS by hand.

Use `pf.feols()` with the IV formula. Hand-computed SEs from two separate regressions are **wrong**.

# Quarter of Birth Exploits Compulsory Schooling to Instrument Education

**Question:** What is the causal return to an additional year of schooling?

**The natural experiment:**

- School entry: must turn 6 by January 1

**Instrument:** Quarter of birth dummies for years of education.

The seasonal pattern exists in high school completion but *not* in college graduation — exactly as compulsory attendance predicts.

**Key estimates (1980 Census, men born 1930–39):**

| | **Return to educ.** |
|---|---|
| OLS | 0.063 |
| | (0.000) |
| Wald (Q1 vs Q3) | 0.089 |
| | (0.033) |
| 2SLS (3 QOB) | 0.076 |
| | (0.029) |

IV $\geq$ OLS $\rightarrow$ consistent with attenuation bias in OLS, not upward ability bias.

# Quarter of Birth Exploits Compulsory Schooling to Instrument Education

**Question:** What is the causal return to an additional year of schooling?

**The natural experiment:**

- School entry: must turn 6 by January 1
- Compulsory attendance: stay until age 16

**Instrument:** Quarter of birth dummies for years of education.

The seasonal pattern exists in high school completion but *not* in college graduation — exactly as compulsory attendance predicts.

**Key estimates (1980 Census, men born 1930–39):**

| | **Return to educ.** |
|---|---|
| OLS | 0.063 |
| | (0.000) |
| Wald (Q1 vs Q3) | 0.089 |
| | (0.033) |
| 2SLS (3 QOB) | 0.076 |
| | (0.029) |

IV $\geq$ OLS $\rightarrow$ consistent with attenuation bias in OLS, not upward ability bias.

# Quarter of Birth Exploits Compulsory Schooling to Instrument Education

**Question:** What is the causal return to an additional year of schooling?

**The natural experiment:**

- School entry: must turn 6 by January 1
- Compulsory attendance: stay until age 16
- Q1 births enter school older $\rightarrow$ reach the dropout age with *fewer* years completed

**Instrument:** Quarter of birth dummies for years of education.

The seasonal pattern exists in high school completion but *not* in college graduation — exactly as compulsory attendance predicts.

**Key estimates (1980 Census, men born 1930–39):**

| | Return to educ. |
|---|---|
| OLS | 0.063 |
| | (0.000) |
| Wald (Q1 vs Q3) | 0.089 |
| | (0.033) |
| 2SLS (3 QOB) | 0.076 |
| | (0.029) |

IV $\geq$ OLS $\rightarrow$ consistent with attenuation bias in OLS, not upward ability bias.

# Weak Instruments: When the First Stage Is Too Small

When $\pi$ is small relative to its standard error, the instrument is weak:

**Consequences of weak instruments:**

- IV estimate biased towards OLS

**The Staiger–Stock rule of thumb:**

- First-stage $F < 10$ signals a weak instrument
- State of the art: Montiel Olea & Pflueger (2015)

**But don't panic (Angrist & Kolesăr 2022):**

In the just-identified case (one instrument, one treatment), the SE increase from weak instruments tends to **cover up the bias**.

You are unlikely to reject $\beta = 0$ spuriously when $F$ is low.

The real danger is in the *over*-identified case with many weak instruments.

# Weak Instruments: When the First Stage Is Too Small

When $\pi$ is small relative to its standard error, the instrument is weak:

**Consequences of weak instruments:**

- IV estimate biased towards OLS
- Standard errors become large (often *very* large)

**The Staiger–Stock rule of thumb:**

- First-stage $F < 10$ signals a weak instrument
- State of the art: Montiel Olea & Pflueger (2015)

**But don't panic (Angrist & Kolesăr 2022):**

In the just-identified case (one instrument, one treatment), the SE increase from weak instruments tends to **cover up the bias**.

You are unlikely to reject $\beta = 0$ spuriously when $F$ is low.

The real danger is in the *over*-identified case with many weak instruments.

# Weak Instruments: When the First Stage Is Too Small

When $\pi$ is small relative to its standard error, the instrument is weak:

**Consequences of weak instruments:**

- IV estimate biased towards OLS
- Standard errors become large (often *very* large)
- Confidence intervals may have poor coverage

**The Staiger–Stock rule of thumb:**

- First-stage $F < 10$ signals a weak instrument
- State of the art: Montiel Olea & Pflueger (2015)

**But don't panic (Angrist & Kolesăr 2022):**

In the just-identified case (one instrument, one treatment), the SE increase from weak instruments tends to **cover up the bias**.

You are unlikely to reject $\beta = 0$ spuriously when $F$ is low.

The real danger is in the *over*-identified case with many weak instruments.

# Just-ID IV Is More Robust Than You Think

Angrist & Kolesăr (2022) show that the reliability of just-identified IV depends on two parameters:

| Parameter | What it measures |
|---|---|
| $E[F]$ | Instrument strength (the population first-stage $F$-statistic) |
| $\rho$ | Endogeneity (correlation between structural and first-stage errors) |

When $\rho < 0.76$, coverage distortion stays below 5 percentage points — even with weak instruments.

**Sign-screening halves the bias:**

Conditioning on the first-stage estimate having the expected sign roughly halves median bias.

In three canonical applications (draft lottery, quarter of birth, same-sex sibship), all calibrated $\rho$ values are moderate.

**Practical advice:** Always check the first-stage $F$-statistic. With one instrument and one endogenous variable, standard IV inference is usually reliable in practice.

# What Does IV Identify?

# Imbens & Angrist (1994): IV Identifies a Local Average Treatment Effect

With heterogeneous treatment effects, IV does *not* identify the average treatment effect for the whole population. It identifies the LATE.

# Imbens & Angrist (1994): IV Identifies a Local Average Treatment Effect

With heterogeneous treatment effects, IV does *not* identify the average treatment effect for the whole population. It identifies the LATE.

**Four types of individuals:**

| Type | $D(Z{=}0)$ | $D(Z{=}1)$ |
|---|---|---|
| Complier | 0 | 1 |
| Always-taker | 1 | 1 |
| Never-taker | 0 | 0 |
| Defier | 1 | 0 |

Under monotonicity ($D_i(1) \geq D_i(0)$ for all $i$), there are no defiers.

**The LATE result:**

$$\beta^{IV} = E[Y_i(1) - Y_i(0) \mid \text{complier}]$$

IV identifies the treatment effect for **compliers** only — those whose treatment status is shifted by the instrument.

Never-takers and always-takers are unaffected by $Z$, so IV is silent about them.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

1. **Internal vs. external validity**

- Your IV estimate applies to *compliers*, not the general population

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

1. **Internal vs. external validity**
   - Your IV estimate applies to *compliers*, not the general population
   - Who are the compliers? Depends on the instrument

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

1. **Internal vs. external validity**

   - Your IV estimate applies to *compliers*, not the general population
   - Who are the compliers? Depends on the instrument
   - Draft lottery compliers $\neq$ quarter-of-birth compliers

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

1. **Internal vs. external validity**
    - Your IV estimate applies to *compliers*, not the general population
    - Who are the compliers? Depends on the instrument
    - Draft lottery compliers $\neq$ quarter-of-birth compliers
    - Two valid instruments can give *different* estimates — and both can be correct

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

### 1. Internal vs. external validity

- Your IV estimate applies to *compliers*, not the general population
- Who are the compliers? Depends on the instrument
- Draft lottery compliers $\neq$ quarter-of-birth compliers
- Two valid instruments can give *different* estimates — and both can be correct

### 2. Monotonicity matters

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

### 1. Internal vs. external validity

- Your IV estimate applies to *compliers*, not the general population
- Who are the compliers? Depends on the instrument
- Draft lottery compliers $\neq$ quarter-of-birth compliers
- Two valid instruments can give *different* estimates — and both can be correct

### 2. Monotonicity matters

- Without monotonicity, IV can have the *wrong sign* even when the true effect is positive for everyone

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

## 1. Internal vs. external validity

- Your IV estimate applies to *compliers*, not the general population
- Who are the compliers? Depends on the instrument
- Draft lottery compliers $\neq$ quarter-of-birth compliers
- Two valid instruments can give *different* estimates — and both can be correct

## 2. Monotonicity matters

- Without monotonicity, IV can have the *wrong sign* even when the true effect is positive for everyone
- Sensible in many settings: draft lottery, RCTs with noncompliance

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# What LATE Means for Your Dissertation

Two practical implications of the LATE result:

### 1. Internal vs. external validity

- Your IV estimate applies to *compliers*, not the general population
- Who are the compliers? Depends on the instrument
- Draft lottery compliers $\neq$ quarter-of-birth compliers
- Two valid instruments can give *different* estimates — and both can be correct

### 2. Monotonicity matters

- Without monotonicity, IV can have the *wrong sign* even when the true effect is positive for everyone
- Sensible in many settings: draft lottery, RCTs with noncompliance
- Potentially questionable in judge IV designs (judges may have different criteria)

**Always discuss who your compliers are.** This is as important as reporting the coefficient.

# Running IV in Python

# Stata to Python: IV Regression Syntax

## Stata

```
ivreg2 y x1 (x_endo = z1 z2),
    first cluster(firm)
```

## Python (pyfixest)

```
pf.feols(
    "y ~ x1 | x_endo ~ z1 + z2",
    vcov={"CRV1": "firm"},
    data=df,
)
```

**The formula has two parts** (without FE) **or three parts** (with FE), separated by |:

| Part | Syntax | Meaning |
|---|---|---|
| Outcome + exogenous | `"y ~ x1"` | Dependent variable and controls |
| Fixed effects (optional) | `"\| fe1 + fe2"` | What to absorb |
| Endogenous ∼ instruments | `"\| x_endo ~ z1 + z2"` | What is instrumented and by what |

With fixed effects: `"y ~ x1 | fe1 + fe2 | x_endo ~ z1 + z2"` — three parts, two pipes.

# Example: IV Regression with Fixed Effects

```python
import pyfixest as pf

# OLS baseline
ols = pf.feols("y ~ treatment + controls | firm + year",
               vcov={"CRV1": "firm"}, data=df)

# IV with instrument
iv = pf.feols("y ~ controls | firm + year | treatment ~ instrument",
              vcov={"CRV1": "firm"}, data=df)

# Compare both models side by side
pf.etable([ols, iv], labels={"treatment": "Treatment",
                             "controls": "Controls"})
```

| Part of formula | What it does |
|---|---|
| `"y ~ controls"` | Dependent variable and exogenous regressors |
| `"\| firm + year"` | Fixed effects to absorb |
| `"\| treatment ~ instrument"` | Endogenous variable instrumented by $Z$ |

# A Weak First Stage Invalidates Your IV Inference

```
# Run IV regression
iv = pf.feols("y ~ controls | firm + year | treatment ~ instrument",
              vcov={"CRV1": "firm"}, data=df)

# View the full summary (includes first-stage F)
iv.summary()

# Access the first-stage F-statistic
print(f"First-stage F: {iv._f_stat_1st_stage:.1f}")
```

| F-statistic | Diagnosis | Action |
|---|---|---|
| $F > 10$ | Adequate strength | Proceed with confidence |
| $F \in [5, 10]$ | Borderline | Report Anderson–Rubin CI |
| $F < 5$ | Weak instrument | Reconsider your instrument |

# Readers Need OLS and IV Side by Side to Judge the Correction

```python
# Build a table comparing OLS and IV
pf.etable(
    [ols, iv],
    labels={"treatment": "Treatment variable",
            "controls": "Control variable"},
    type="tex",  # export for your dissertation
)
```

**What every IV table in your dissertation should include:**

1. OLS estimate alongside the IV estimate

# Readers Need OLS and IV Side by Side to Judge the Correction

```python
# Build a table comparing OLS and IV
pf.etable(
    [ols, iv],
    labels={"treatment": "Treatment variable",
            "controls": "Control variable"},
    type="tex",  # export for your dissertation
)
```

**What every IV table in your dissertation should include:**

1. OLS estimate alongside the IV estimate
2. First-stage coefficient and $F$-statistic

# Readers Need OLS and IV Side by Side to Judge the Correction

```python
# Build a table comparing OLS and IV
pf.etable(
    [ols, iv],
    labels={"treatment": "Treatment variable",
            "controls": "Control variable"},
    type="tex",  # export for your dissertation
)
```

**What every IV table in your dissertation should include:**

1. OLS estimate alongside the IV estimate
2. First-stage coefficient and *F*-statistic
3. The reduced-form estimate (optional but encouraged)

# Readers Need OLS and IV Side by Side to Judge the Correction

```python
# Build a table comparing OLS and IV
pf.etable(
    [ols, iv],
    labels={"treatment": "Treatment variable",
            "controls": "Control variable"},
    type="tex",  # export for your dissertation
)
```

**What every IV table in your dissertation should include:**

1. OLS estimate alongside the IV estimate
2. First-stage coefficient and *F*-statistic
3. The reduced-form estimate (optional but encouraged)
4. Standard errors clustered at the appropriate level

# Readers Need OLS and IV Side by Side to Judge the Correction

```python
# Build a table comparing OLS and IV
pf.etable(
    [ols, iv],
    labels={"treatment": "Treatment variable",
            "controls": "Control variable"},
    type="tex",  # export for your dissertation
)
```

**What every IV table in your dissertation should include:**

1. OLS estimate alongside the IV estimate
2. First-stage coefficient and *F*-statistic
3. The reduced-form estimate (optional but encouraged)
4. Standard errors clustered at the appropriate level
5. A note describing the instrument

# OLS Overstates the Effect by 60%; IV Cuts It Back to the Truth

Simulated data: true $\beta = 0.8$, but $x$ is endogenous ($\rho = 0.5$). Instrument $z$ is relevant ($F > 950$) and valid.

| | (1) OLS | (2) IV | (3) IV + FE |
|---|---|---|---|
| Treatment ($x$) | 1.291*** | 0.744*** | 0.744*** |
| | (0.012) | (0.036) | (0.037) |
| Control ($w$) | 0.493*** | 0.488*** | 0.488*** |
| | (0.012) | (0.015) | (0.014) |
| Fixed effects | — | — | Firm, Year |
| First-stage $F$ | — | 954.5 | 1099.4 |
| Observations | 5,000 | 5,000 | 5,000 |

# OLS Overstates the Effect by 60%; IV Cuts It Back to the Truth

Simulated data: true $\beta = 0.8$, but $x$ is endogenous ($\rho = 0.5$). Instrument $z$ is relevant ($F > 950$) and valid.

| | (1) OLS | (2) IV | (3) IV + FE |
|---|---|---|---|
| Treatment ($x$) | 1.291*** | 0.744*** | 0.744*** |
| | (0.012) | (0.036) | (0.037) |
| Control ($w$) | 0.493*** | 0.488*** | 0.488*** |
| | (0.012) | (0.015) | (0.014) |
| Fixed effects | — | — | Firm, Year |
| First-stage $F$ | — | 954.5 | 1099.4 |
| Observations | 5,000 | 5,000 | 5,000 |

**What to notice:**

- OLS: $\hat{\beta} = 1.291$ — biased upward from 0.8 because $\mathrm{Corr}(x, \varepsilon) > 0$

# OLS Overstates the Effect by 60%; IV Cuts It Back to the Truth

Simulated data: true $\beta = 0.8$, but $x$ is endogenous ($\rho = 0.5$). Instrument $z$ is relevant ($F > 950$) and valid.

| | (1) OLS | (2) IV | (3) IV + FE |
|---|---|---|---|
| Treatment ($x$) | 1.291*** | 0.744*** | 0.744*** |
| | (0.012) | (0.036) | (0.037) |
| Control ($w$) | 0.493*** | 0.488*** | 0.488*** |
| | (0.012) | (0.015) | (0.014) |
| Fixed effects | — | — | Firm, Year |
| First-stage $F$ | — | 954.5 | 1099.4 |
| Observations | 5,000 | 5,000 | 5,000 |

**What to notice:**

- OLS: $\hat{\beta} = 1.291$ — biased upward from 0.8 because Corr$(x, \varepsilon) > 0$
- IV: $\hat{\beta} = 0.744$ — close to the true value; larger SE is the price of consistency

# OLS Overstates the Effect by 60%; IV Cuts It Back to the Truth

Simulated data: true $\beta = 0.8$, but $x$ is endogenous ($\rho = 0.5$). Instrument $z$ is relevant ($F > 950$) and valid.

| | (1) OLS | (2) IV | (3) IV + FE |
|---|---|---|---|
| Treatment ($x$) | 1.291*** | 0.744*** | 0.744*** |
| | (0.012) | (0.036) | (0.037) |
| Control ($w$) | 0.493*** | 0.488*** | 0.488*** |
| | (0.012) | (0.015) | (0.014) |
| Fixed effects | — | — | Firm, Year |
| First-stage $F$ | — | 954.5 | 1099.4 |
| Observations | 5,000 | 5,000 | 5,000 |

**What to notice:**

- OLS: $\hat{\beta} = 1.291$ — biased upward from 0.8 because $\text{Corr}(x, \varepsilon) > 0$
- IV: $\hat{\beta} = 0.744$ — close to the true value; larger SE is the price of consistency
- IV + FE: same coefficient, confirming FE alone cannot fix endogeneity

Simulated data: true $\beta = 0.8$, but $x$ is endogenous ($\rho = 0.5$). Instrument $z$ is relevant ($F > 950$) and valid.

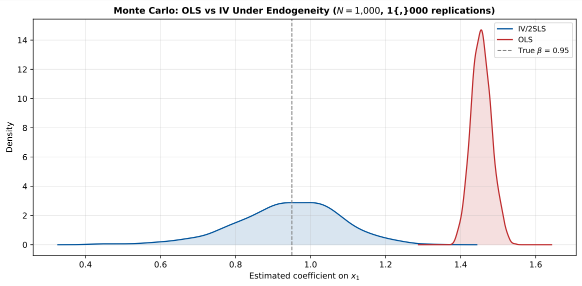
| | (1) OLS | (2) IV | (3) IV + FE |
|---|---|---|---|
| Treatment ($x$) | 1.291*** | 0.744*** | 0.744*** |
| | (0.012) | (0.036) | (0.037) |
| Control ($w$) | 0.493*** | 0.488*** | 0.488*** |
| | (0.012) | (0.015) | (0.014) |
| Fixed effects | — | — | Firm, Year |
| First-stage $F$ | — | 954.5 | 1099.4 |
| Observations | 5,000 | 5,000 | 5,000 |

**What to notice:**

- OLS: $\hat{\beta} = 1.291$ — biased upward from 0.8 because $\text{Corr}(x, \varepsilon) > 0$
- IV: $\hat{\beta} = 0.744$ — close to the true value; larger SE is the price of consistency
- IV + FE: same coefficient, confirming FE alone cannot fix endogeneity
- First-stage $F \gg 10$ — no weak instrument concern

# Simulation Confirms: OLS Is Biased, IV Recovers the Truth

When $x_1$ is endogenous ($\rho = 0.5$), OLS is biased upward. IV, using a valid instrument, centres on the true $\beta = 0.95$:




| | OLS | IV/2SLS |
|---|---|---|
| Mean | 1.4554 | 0.9414 |
| Std | 0.0264 | 0.1373 |
| Min | 1.3876 | 0.4263 |
| Q1 | 1.4368 | 0.8600 |
| Median | 1.4544 | 0.9476 |
| Q3 | 1.4734 | 1.0352 |
| Max | 1.5422 | 1.3428 |
| True $\beta$ | 0.95 | |
| Bias | +0.5054 | -0.0086 |

1,000 replications, $N = 1{,}000$ per sample.

# Advanced: Shift-Share Instruments

Many instruments are constructed from multiple sources of variation:

$$Z_i = \sum_k s_{ik} \, g_k$$

where $g_k$ are shocks and $s_{ik}$ are exposure shares.

**Example: Autor, Dorn & Hanson (2013)**

- $g_k$: Chinese import growth in industry $k$ (measured in other countries)
- $s_{ik}$: share of industry $k$ in commuting zone $i$'s employment
- $Z_i$: predicted Chinese import exposure for CZ $i$

**BHJ key insight:**

Identification can come from quasi-random *shocks* $g_k$, even if the *shares* $s_{ik}$ are endogenous.

This requires: (1) many independent shocks, (2) controlling for $\sum_k s_{ik} q_k$ to isolate as-good-as-random variation.

# Event Studies

# Event Studies Measure the Market's Reaction to News

An event study measures how a specific event — an earnings announcement, a regulatory sanction, a cyberattack — affects firm value.

**The classical approach (this week):**

- Estimate a "normal return" benchmark

Best for: M&A, earnings surprises, regulatory actions, cyberattacks — any event with a precise date and a stock-price outcome.

**Why event studies work:**

Under semi-strong efficiency, stock prices adjust quickly to new information. The abnormal return on the event day captures the market's *valuation* of the news.

**Next week** we cover the *panel* (DiD-style) event study — staggered treatments, parallel trends, and dynamic effects.

# Event Studies Measure the Market's Reaction to News

An event study measures how a specific event — an earnings announcement, a regulatory sanction, a cyberattack — affects firm value.

**The classical approach (this week):**

- Estimate a "normal return" benchmark
- Compute abnormal returns (AR) during the event window

Best for: M&A, earnings surprises, regulatory actions, cyberattacks — any event with a precise date and a stock-price outcome.

**Why event studies work:**

Under semi-strong efficiency, stock prices adjust quickly to new information. The abnormal return on the event day captures the market's *valuation* of the news.

**Next week** we cover the *panel* (DiD-style) event study — staggered treatments, parallel trends, and dynamic effects.

# Event Studies Measure the Market's Reaction to News

An event study measures how a specific event — an earnings announcement, a regulatory sanction, a cyberattack — affects firm value.

**The classical approach (this week):**

- Estimate a "normal return" benchmark
- Compute abnormal returns (AR) during the event window
- Cumulate and average across firms (CAR)

**Why event studies work:**

Under semi-strong efficiency, stock prices adjust quickly to new information. The abnormal return on the event day captures the market's *valuation* of the news.

Best for: M&A, earnings surprises, regulatory actions, cyberattacks — any event with a precise date and a stock-price outcome.

**Next week** we cover the *panel* (DiD-style) event study — staggered treatments, parallel trends, and dynamic effects.

# Event Studies Measure the Market's Reaction to News

An event study measures how a specific event — an earnings announcement, a regulatory sanction, a cyberattack — affects firm value.

**The classical approach (this week):**

- Estimate a "normal return" benchmark
- Compute abnormal returns (AR) during the event window
- Cumulate and average across firms (CAR)
- Test whether the average effect is statistically significant

Best for: M&A, earnings surprises, regulatory actions, cyberattacks — any event with a precise date and a stock-price outcome.

**Why event studies work:**

Under semi-strong efficiency, stock prices adjust quickly to new information. The abnormal return on the event day captures the market's *valuation* of the news.

**Next week** we cover the *panel* (DiD-style) event study — staggered treatments, parallel trends, and dynamic effects.

# MacKinlay (1997) Codified the Five-Step Event Study Recipe

1. **Define the event** and collect event dates

The logic is simple: if the event had no effect, the stock should have earned its "normal" return. Any deviation is the **abnormal return**.

# MacKinlay (1997) Codified the Five-Step Event Study Recipe

1. **Define the event** and collect event dates
2. **Choose windows:** estimation window (e.g., $[-250, -11]$) and event window (e.g., $[-10, +10]$)

The logic is simple: if the event had no effect, the stock should have earned its "normal" return. Any deviation is the **abnormal return**.

# MacKinlay (1997) Codified the Five-Step Event Study Recipe

1. **Define the event** and collect event dates
2. **Choose windows:** estimation window (e.g., $[-250, -11]$) and event window (e.g., $[-10, +10]$)
3. **Estimate normal returns** using a benchmark model (market model, Fama–French)

The logic is simple: if the event had no effect, the stock should have earned its "normal" return. Any deviation is the **abnormal return**.

# MacKinlay (1997) Codified the Five-Step Event Study Recipe

1. **Define the event** and collect event dates
2. **Choose windows:** estimation window (e.g., $[-250, -11]$) and event window (e.g., $[-10, +10]$)
3. **Estimate normal returns** using a benchmark model (market model, Fama–French)
4. **Compute abnormal returns:** $AR_{i,t} = R_{i,t} - \hat{R}_{i,t}$

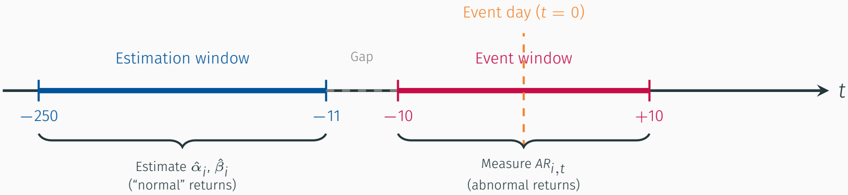
The logic is simple: if the event had no effect, the stock should have earned its "normal" return. Any deviation is the **abnormal return**.

# MacKinlay (1997) Codified the Five-Step Event Study Recipe

1. **Define the event** and collect event dates
2. **Choose windows:** estimation window (e.g., $[-250, -11]$) and event window (e.g., $[-10, +10]$)
3. **Estimate normal returns** using a benchmark model (market model, Fama–French)
4. **Compute abnormal returns:** $AR_{i,t} = R_{i,t} - \hat{R}_{i,t}$
5. **Aggregate and test:** cumulate over time (CAR), average across firms ($\overline{CAR}$), test $H_0$: no effect

The logic is simple: if the event had no effect, the stock should have earned its "normal" return. Any deviation is the **abnormal return**.

# The Estimation Window Must Not Overlap the Event Window




If the estimation window includes the event, the market model absorbs the effect you are trying to measure. A gap of $\sim$10 trading days is standard.

# The Market Model Is the Workhorse for Normal Returns

The market model relates each stock's return to the market:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}$$

Estimate $\hat{\alpha}_i$ and $\hat{\beta}_i$ by OLS over the estimation window ($\sim$240 trading days). Then the abnormal return is:

$$AR_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t}$$

**Why the market model?**

- Removes market-wide movements

**Alternative benchmarks:**

| Model | When to use |
|---|---|
| Market model | Default choice |
| FF 3-factor | Higher $R^2$ |
| Mean return | Simple; similar power |

# The Market Model Is the Workhorse for Normal Returns

The market model relates each stock's return to the market:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}$$

Estimate $\hat{\alpha}_i$ and $\hat{\beta}_i$ by OLS over the estimation window (∼240 trading days). Then the abnormal return is:

$$AR_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t}$$

**Why the market model?**

- Removes market-wide movements
- Higher $R^2$ → lower variance of $AR$

**Alternative benchmarks:**

| Model | When to use |
|---|---|
| Market model | Default choice |
| FF 3-factor | Higher $R^2$ |
| Mean return | Simple; similar power |

# The Market Model Is the Workhorse for Normal Returns

The market model relates each stock's return to the market:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}$$

Estimate $\hat{\alpha}_i$ and $\hat{\beta}_i$ by OLS over the estimation window ($\sim$240 trading days). Then the abnormal return is:

$$AR_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t}$$

**Why the market model?**

- Removes market-wide movements
- Higher $R^2 \rightarrow$ lower variance of $AR$
- Brown & Warner (1985): performs as well as more complex alternatives

**Alternative benchmarks:**

| Model | When to use |
|---|---|
| Market model | Default choice |
| FF 3-factor | Higher $R^2$ |
| Mean return | Simple; similar power |

# Aggregation Turns Firm-Level Noise into a Testable Hypothesis

**Within a firm** — cumulate over the event window:

$$CAR_i(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} AR_{i,t}$$

**Across firms** — average the CARs:

$$\overline{CAR}(\tau_1, \tau_2) = \frac{1}{N} \sum_{i=1}^{N} CAR_i(\tau_1, \tau_2)$$

**Test statistic** (under $H_0$: no effect):

$$\theta = \frac{\overline{CAR}}{\sqrt{\mathrm{Var}(\overline{CAR})}} \sim N(0,1)$$

Always report: $\overline{CAR}$, its SE, the $t$-statistic, and $N$. Add a non-parametric test for robustness.

# Daily Data Yield Three Times the Power of Monthly Data

Brown & Warner (1985) simulated 12,500 event studies with actual CRSP returns:

| Design choice | Power |
|---|---|
| Daily, 1-day window | 80% |
| Daily, 11-day window | 13% |
| Monthly returns | 27% |

$N = 50$, abnormal return $= 1\%$, market model.

**Key takeaways:**

- Use daily data whenever possible

**MacKinlay (1997) power table:**

| $N$ | $AR$=0.5% | $AR$=1% |
|---|---|---|
| 20 | 0.20 | 0.61 |
| 50 | 0.42 | 0.94 |
| 100 | 0.71 | 1.00 |
| 200 | 0.94 | 1.00 |

$\sigma = 0.02$, 5% two-sided test.

With a 0.5% abnormal return, you need $N > 60$ events for 50% power.

# Daily Data Yield Three Times the Power of Monthly Data

Brown & Warner (1985) simulated 12,500 event studies with actual CRSP returns:

| Design choice | Power |
|---|---|
| Daily, 1-day window | 80% |
| Daily, 11-day window | 13% |
| Monthly returns | 27% |

$N = 50$, abnormal return = 1%, market model.

**Key takeaways:**

- Use daily data whenever possible
- Keep the event window short

**MacKinlay (1997) power table:**

| $N$ | $AR$=0.5% | $AR$=1% |
|---|---|---|
| 20 | 0.20 | 0.61 |
| 50 | 0.42 | 0.94 |
| 100 | 0.71 | 1.00 |
| 200 | 0.94 | 1.00 |

$\sigma = 0.02$, 5% two-sided test.

With a 0.5% abnormal return, you need $N > 60$ events for 50% power.

# Daily Data Yield Three Times the Power of Monthly Data

Brown & Warner (1985) simulated 12,500 event studies with actual CRSP returns:

| Design choice | Power |
|---|---|
| Daily, 1-day window | 80% |
| Daily, 11-day window | 13% |
| Monthly returns | 27% |

$N = 50$, abnormal return $= 1\%$, market model.

**Key takeaways:**

- Use daily data whenever possible
- Keep the event window short
- OLS market model works as well as Scholes–Williams or Dimson adjustments

**MacKinlay (1997) power table:**

| $N$ | $AR$=0.5% | $AR$=1% |
|---|---|---|
| 20 | 0.20 | 0.61 |
| 50 | 0.42 | 0.94 |
| 100 | 0.71 | 1.00 |
| 200 | 0.94 | 1.00 |

$\sigma = 0.02$, 5% two-sided test.

With a 0.5% abnormal return, you need $N > 60$ events for 50% power.

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)
2. **Confounding events:** Is the event window free from other news?

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)

2. **Confounding events:** Is the event window free from other news?
   → Armour et al. (2017) dropped 43 of 83 cases for leakage or confounding

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   - $\rightarrow$ Imprecise dates destroy power (Brown & Warner 1985)
2. **Confounding events:** Is the event window free from other news?
   - $\rightarrow$ Armour et al. (2017) dropped 43 of 83 cases for leakage or confounding
3. **Sample size:** Do you have enough events for statistical power?

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)
2. **Confounding events:** Is the event window free from other news?
   → Armour et al. (2017) dropped 43 of 83 cases for leakage or confounding
3. **Sample size:** Do you have enough events for statistical power?
   → Use MacKinlay's power table to check *before* you start

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)
2. **Confounding events:** Is the event window free from other news?
   → Armour et al. (2017) dropped 43 of 83 cases for leakage or confounding
3. **Sample size:** Do you have enough events for statistical power?
   → Use MacKinlay's power table to check *before* you start
4. **Clustering:** Do many events fall on the same calendar day?

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)
2. **Confounding events:** Is the event window free from other news?
   → Armour et al. (2017) dropped 43 of 83 cases for leakage or confounding
3. **Sample size:** Do you have enough events for statistical power?
   → Use MacKinlay's power table to check *before* you start
4. **Clustering:** Do many events fall on the same calendar day?
   → If yes, use cross-sectional variance estimators (Boehmer et al. 1991)

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)
2. **Confounding events:** Is the event window free from other news?
   → Armour et al. (2017) dropped 43 of 83 cases for leakage or confounding
3. **Sample size:** Do you have enough events for statistical power?
   → Use MacKinlay's power table to check *before* you start
4. **Clustering:** Do many events fall on the same calendar day?
   → If yes, use cross-sectional variance estimators (Boehmer et al. 1991)
5. **Cross-sectional analysis:** What explains variation in CARs?

# Event Study Design Checklist

Before you commit to an event study, work through these questions:

1. **Event dates:** Can you identify the *exact* day the market learns the news?
   → Imprecise dates destroy power (Brown & Warner 1985)
2. **Confounding events:** Is the event window free from other news?
   → Armour et al. (2017) dropped 43 of 83 cases for leakage or confounding
3. **Sample size:** Do you have enough events for statistical power?
   → Use MacKinlay's power table to check *before* you start
4. **Clustering:** Do many events fall on the same calendar day?
   → If yes, use cross-sectional variance estimators (Boehmer et al. 1991)
5. **Cross-sectional analysis:** What explains variation in CARs?
   → Regress CARs on firm/event characteristics with robust SEs

# Regulatory Fines Are Dwarfed by Reputational Losses

Armour, Mayer & Polo (2017) exploit the U.K.'s single-announcement enforcement regime — one press release reveals both the misconduct and the penalty.

# Regulatory Fines Are Dwarfed by Reputational Losses

Armour, Mayer & Polo (2017) exploit the U.K.'s single-announcement enforcement regime — one press release reveals both the misconduct and the penalty.

| | **Customers** | **Third party** |
|---|---|---|
| CAR($-1,+1$) | $-2.62\%^{***}$ | $+0.24\%$ |
| Fine (% mkt cap) | $-0.13\%$ | $-0.19\%$ |
| Reputational loss | $-2.31\%^{**}$ | $+0.43\%$ |
| $N$ | 26 | 14 |

Reputational losses are $\approx 9\times$ the fines — but *only* when customers are harmed.

# Regulatory Fines Are Dwarfed by Reputational Losses

Armour, Mayer & Polo (2017) exploit the U.K.'s single-announcement enforcement regime — one press release reveals both the misconduct and the penalty.

| | Customers | Third party |
|---|---|---|
| CAR(−1,+1) | −2.62%*** | +0.24% |
| Fine (% mkt cap) | −0.13% | −0.19% |
| Reputational loss | −2.31%** | +0.43% |
| *N* | 26 | 14 |

Reputational losses are ≈ 9× the fines — but *only* when customers are harmed.

**Why this paper matters:**

- Clean identification: single announcement, no leakage

# Regulatory Fines Are Dwarfed by Reputational Losses

Armour, Mayer & Polo (2017) exploit the U.K.'s single-announcement enforcement regime — one press release reveals both the misconduct and the penalty.

| | Customers | Third party |
|---|---|---|
| CAR($-1,+1$) | $-2.62\%^{***}$ | $+0.24\%$ |
| Fine (% mkt cap) | $-0.13\%$ | $-0.19\%$ |
| Reputational loss | $-2.31\%^{**}$ | $+0.43\%$ |
| $N$ | 26 | 14 |

Reputational losses are $\approx 9\times$ the fines — but *only* when customers are harmed.

**Why this paper matters:**

- Clean identification: single announcement, no leakage
- Clean decomposition: fine vs. reputation

# Regulatory Fines Are Dwarfed by Reputational Losses

Armour, Mayer & Polo (2017) exploit the U.K.'s single-announcement enforcement regime — one press release reveals both the misconduct and the penalty.

| | Customers | Third party |
|---|---|---|
| CAR(−1,+1) | −2.62%*** | +0.24% |
| Fine (% mkt cap) | −0.13% | −0.19% |
| Reputational loss | −2.31%** | +0.43% |
| $N$ | 26 | 14 |

Reputational losses are ≈ 9× the fines — but *only* when customers are harmed.

**Why this paper matters:**

- Clean identification: single announcement, no leakage
- Clean decomposition: fine vs. reputation
- Third-party wrongs produce *no* market penalty

# Regulatory Fines Are Dwarfed by Reputational Losses

Armour, Mayer & Polo (2017) exploit the U.K.'s single-announcement enforcement regime — one press release reveals both the misconduct and the penalty.

| | Customers | Third party |
|---|---|---|
| CAR($-1,+1$) | $-2.62\%$*** | $+0.24\%$ |
| Fine (% mkt cap) | $-0.13\%$ | $-0.19\%$ |
| Reputational loss | $-2.31\%$** | $+0.43\%$ |
| $N$ | 26 | 14 |

Reputational losses are $\approx 9\times$ the fines — but *only* when customers are harmed.

**Why this paper matters:**

- Clean identification: single announcement, no leakage
- Clean decomposition: fine vs. reputation
- Third-party wrongs produce *no* market penalty
- 40 hand-collected events

# Cyberattacks Cost Shareholders 0.84% in Three Days

Kamiya, Kang, Kim, Milidonis & Stulz (2021) study 165 cyberattacks on listed U.S. firms (2005–2017):

| Measure | Estimate |
|---|---|
| CAR(−1,+1) | −0.84%*** |
| Share: out-of-pocket costs | ≈ 1–7% |
| Share: reputation loss | ≈ 93–99% |
| Post-attack sales growth | −3.4 pp |
| Board risk oversight | +19 pp |

93–99% of shareholder losses are reputational. Firms respond by improving governance.

**Design features:**

- Market model + FF3 + FF4 benchmarks
- 220-day estimation window (−280 to −61)
- DiD with propensity-score matching for post-attack outcomes
- Cross-sectional regressions of CARs on attack characteristics

# Build the Classical Event Study Pipeline in Python

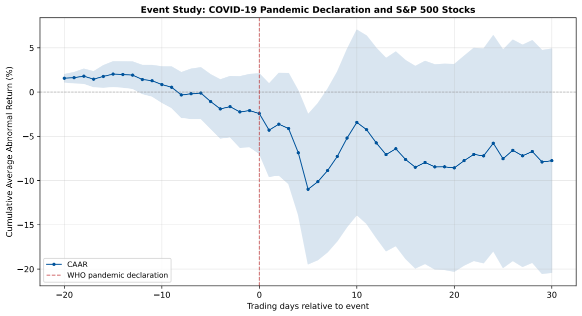
```python
import polars as pl
import statsmodels.formula.api as smf

# 1. Load returns and merge with market returns
returns = pl.read_csv("daily_returns.csv", try_parse_dates=True)
market = pl.read_csv("market_returns.csv", try_parse_dates=True)
df = returns.join(market, on="date")
# 2. For each event, split into estimation and event windows
est_start, est_end = -250, -11   # estimation window
evt_start, evt_end = -10, 10     # event window

# 3. Estimate market model over estimation window
est_data = df.filter(
    (pl.col("rel_day") >= est_start) & (pl.col("rel_day") <= est_end)
)
model = smf.ols("ret ~ mkt_ret", data=est_data.to_pandas()).fit()

# 4. Predict normal returns and compute abnormal returns
evt_data = df.filter(
    (pl.col("rel_day") >= evt_start) & (pl.col("rel_day") <= evt_end)
)
evt_pd = evt_data.to_pandas()
evt_pd["ar"] = evt_pd["ret"] - model.predict(evt_pd)
```

# Compute and Plot Cumulative Abnormal Returns

```python
import matplotlib.pyplot as plt

# Average abnormal returns across all events
aar = evt_pd.groupby("rel_day")["ar"].mean()

# Cumulative average abnormal returns
caar = aar.cumsum()

# Plot
fig, ax = plt.subplots(figsize=(10, 5))
ax.plot(caar.index, caar.values, "o-", color="#00539b")
ax.axhline(0, color="gray", linestyle="--", linewidth=0.8)
ax.axvline(0, color="#be2a2c", linestyle="--", alpha=0.7,
           label="Event day")
ax.set_xlabel("Days relative to event")
ax.set_ylabel("Cumulative Average AR (%)")
plt.savefig("event_study_car.pdf", bbox_inches="tight")
```

Flat pre-event CARs support your identification. A jump at $t = 0$ is your treatment effect. Post-event drift suggests incomplete adjustment.

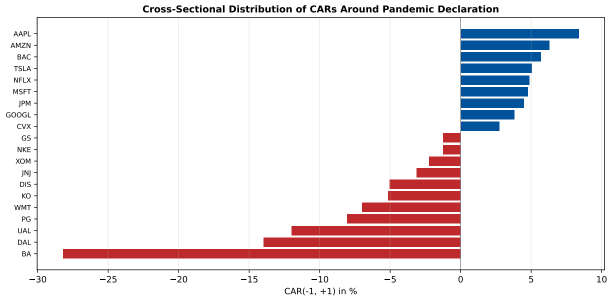
# COVID-19 Pandemic Declaration: Real CAAR from 20 S&P 500 Stocks




| Window | Mean CAR | $t$-stat | Wilcoxon $p$ | $N$ |
|---|---|---|---|---|
| CAR(-1,1) | -2.06% | -1.04 | 0.5459 | 20 |
| CAR(-5,5) | -10.87% | -1.62 | 0.3488 | 20 |
| CAR(-1,10) | -1.18% | -0.45 | 0.9273 | 20 |
| CAR(0,0) | -0.34% | -0.35 | 0.6742 | 20 |

Market model estimated over 2019 (251 days). Event: WHO pandemic declaration, 11 March 2020. 20 large-cap stocks across sectors.

# Cross-Sectional CARs Reveal Sector Heterogeneity




## Daily AAR around the event:

| Day | AAR (%) | CAAR (%) | *t*-stat | *N* |
|---|---|---|---|---|
| -5 | -0.95* | -1.07 | -1.91 | 20 |
| -4 | -0.84 | -1.91 | -1.28 | 20 |
| -3 | 0.26 | -1.65 | 0.57 | 20 |
| -2 | -0.60 | -2.25 | -0.57 | 20 |
| -1 | 0.15 | -2.09 | 0.33 | 20 |
| +0 | -0.34 | -2.43 | -0.35 | 20 |
| +1 | -1.87 | -4.30 | -1.38 | 20 |
| +2 | 0.67 | -3.63 | 0.55 | 20 |
| +3 | -0.49 | -4.12 | -0.40 | 20 |
| +4 | -2.74* | -6.86 | -1.70 | 20 |
| +5 | -4.12* | -10.98 | -1.70 | 20 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1,+1)$ and $(-5,+5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)
3. The number of events $N$ and the estimation window length

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)
3. The number of events $N$ and the estimation window length
4. A CAAR plot with confidence bands and the event day marked

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)
3. The number of events $N$ and the estimation window length
4. A CAAR plot with confidence bands and the event day marked

**Common pitfalls:**

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)
3. The number of events $N$ and the estimation window length
4. A CAAR plot with confidence bands and the event day marked

**Common pitfalls:**

- Reporting only one event window — reviewers will ask for others

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)
3. The number of events $N$ and the estimation window length
4. A CAAR plot with confidence bands and the event day marked

**Common pitfalls:**

- Reporting only one event window — reviewers will ask for others
- Omitting the market model diagnostics ($R^2$, number of estimation days)

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)
3. The number of events $N$ and the estimation window length
4. A CAAR plot with confidence bands and the event day marked

**Common pitfalls:**

- Reporting only one event window — reviewers will ask for others
- Omitting the market model diagnostics ($R^2$, number of estimation days)
- Using too long an event window, which dilutes the signal with noise

# Presenting Event Study Results in Your Dissertation

**Every event study should report:**

1. $\overline{CAR}$ for at least two windows (e.g., $(-1, +1)$ and $(-5, +5)$)
2. A $t$-test and a non-parametric test (e.g., Wilcoxon signed-rank)
3. The number of events $N$ and the estimation window length
4. A CAAR plot with confidence bands and the event day marked

**Common pitfalls:**

- Reporting only one event window — reviewers will ask for others
- Omitting the market model diagnostics ($R^2$, number of estimation days)
- Using too long an event window, which dilutes the signal with noise
- Not discussing pre-event CARs — a drift before the event suggests information leakage

# Summary

# Key Papers Covered Today

| Paper | Key contribution |
|---|---|
| *Instrumental variables* | |
| Angrist (1990) | Draft lottery IV; white veterans lost ≈15% in earnings. |
| Imbens & Angrist (1994) | IV identifies the LATE for compliers, not the whole population. |
| Angrist & Krueger (1991) | Quarter of birth as IV for schooling; canonical 2SLS example. |
| Angrist & Kolesăr (2022) | Just-identified IV is more robust to weak instruments than previously thought. |
| *Event studies* | |
| MacKinlay (1997) | Five-step recipe, power tables, cross-sectional tests. |
| Armour et al. (2017) | Reputational losses ≈ 9× fines in U.K. enforcement. |
| Kamiya et al. (2021) | 93–99% of cyberattack losses are reputational. |

# IV Checklist for Your Dissertation

Before you commit to an IV strategy, work through these questions:

1. **What is your instrument?** Can you state the source of exogenous variation in one sentence?

# IV Checklist for Your Dissertation

Before you commit to an IV strategy, work through these questions:

1. **What is your instrument?** Can you state the source of exogenous variation in one sentence?
2. **Relevance:** Is the first-stage $F$-statistic above 10?

# IV Checklist for Your Dissertation

Before you commit to an IV strategy, work through these questions:

1. **What is your instrument?** Can you state the source of exogenous variation in one sentence?
2. **Relevance:** Is the first-stage $F$-statistic above 10?
3. **Independence:** Why is $Z$ uncorrelated with confounders? Can you test balance on observables?

# IV Checklist for Your Dissertation

Before you commit to an IV strategy, work through these questions:

1. **What is your instrument?** Can you state the source of exogenous variation in one sentence?
2. **Relevance:** Is the first-stage $F$-statistic above 10?
3. **Independence:** Why is $Z$ uncorrelated with confounders? Can you test balance on observables?
4. **Exclusion:** Does $Z$ affect $Y$ only through $X$? What alternative channels could there be?

# IV Checklist for Your Dissertation

Before you commit to an IV strategy, work through these questions:

1. **What is your instrument?** Can you state the source of exogenous variation in one sentence?
2. **Relevance:** Is the first-stage $F$-statistic above 10?
3. **Independence:** Why is $Z$ uncorrelated with confounders? Can you test balance on observables?
4. **Exclusion:** Does $Z$ affect $Y$ only through $X$? What alternative channels could there be?
5. **Who are the compliers?** Can you characterise the sub-population whose treatment effect you estimate?

# IV Checklist for Your Dissertation

Before you commit to an IV strategy, work through these questions:

1. **What is your instrument?** Can you state the source of exogenous variation in one sentence?
2. **Relevance:** Is the first-stage $F$-statistic above 10?
3. **Independence:** Why is $Z$ uncorrelated with confounders? Can you test balance on observables?
4. **Exclusion:** Does $Z$ affect $Y$ only through $X$? What alternative channels could there be?
5. **Who are the compliers?** Can you characterise the sub-population whose treatment effect you estimate?
6. **Do you report the reduced form and first stage?** Both regressions should appear in your paper.

# Week 6 — Key Takeaways

**Instrumental variables:**

- IV solves time-varying confounders and reverse causality that FE cannot
- Three conditions: relevance, independence, exclusion
- Always check the first-stage $F > 10$
- IV identifies the LATE for compliers
- In Python: `pf.feols("y ~ x | fe | endo ~ z")`

**Classical event studies:**

- Market model $\rightarrow$ abnormal returns $\rightarrow$ CAR
- Use daily data and keep the event window short
- Report multiple CAR windows with parametric and non-parametric tests
- Check pre-event CARs for leakage
- Cross-sectional regressions explain *why* CARs vary

**Next week:** Difference-in-differences, panel event studies, and regression discontinuity.