A causal approach to test empirical capital structure regularities

Simone Cenci[1] and Stephen Kealhofer[2]

[1]*Imperial College Business School, Exhibition Rd, London SW7 2BX*

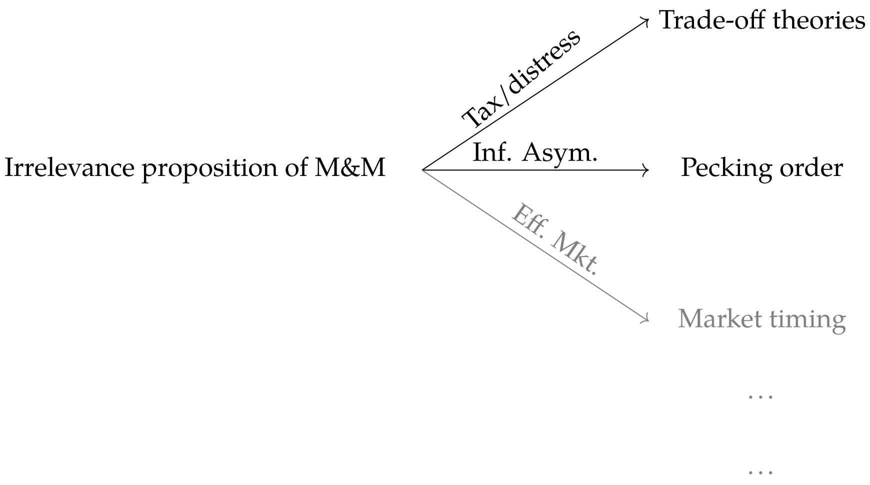
[2]*Blackstone Credit, San Francisco, 101 California St., CA, 94111, USA*

January 31st, 2023

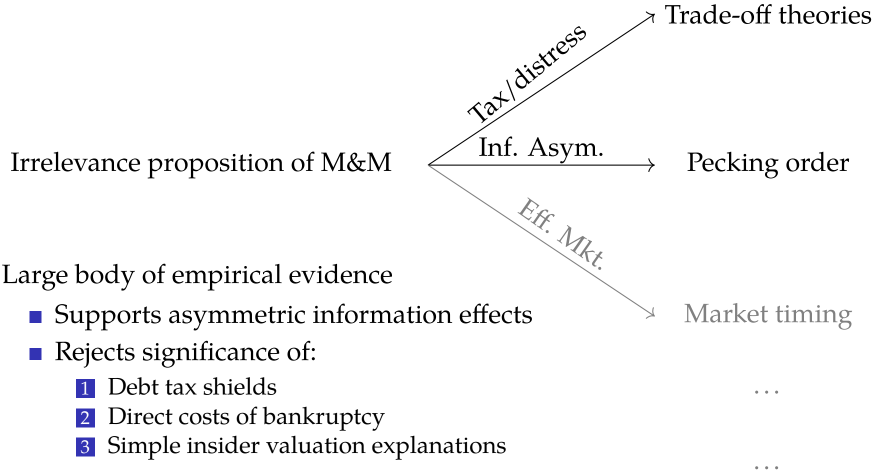
One-slider on capital structure theories

Irrelevance proposition of M&M

One-slider on capital structure theories



One-slider on capital structure theories



Large body of empirical evidence

- Supports asymmetric information effects
- Rejects significance of:
  1. Debt tax shields
  2. Direct costs of bankruptcy
  3. Simple insider valuation explanations

One-slider on capital structure theories

Irrelevance proposition of M&M

Tax/distress → Trade-off theories

Inf. Asym. → Pecking order

Eff. Mkt. → Market timing

...

...

Why is discussion still framed by these theories?

- Simple narratives
- There must be a trade-off somewhere here
- Asymmetric information matters
- Timing shows up empirically

Contemporaneous correlation model

A set of empirical findings, initially summarized by Zingales (1995) and subsequently more exhaustively by Frank and Goyal (2009)

- Consistent historic and cross-border evidence of leverage related to certain firm characteristics

Contemporaneous correlation model

A set of empirical findings, initially summarized by Zingales (1995) and subsequently more exhaustively by Frank and Goyal (2009)

- Consistent historic and cross-border evidence of leverage related to certain firm characteristics
- Depending on the interpretation of these firm characteristics, they play into virtually all the applied modeling papers

Contemporaneous correlation model

A set of empirical findings, initially summarized by Zingales (1995) and subsequently more exhaustively by Frank and Goyal (2009)

- Consistent historic and cross-border evidence of leverage related to certain firm characteristics
- Depending on the interpretation of these firm characteristics, they play into virtually all the applied modeling papers
- Because the interpretations are different, can be used to support quite different theories

Contemporaneous correlation model

A set of empirical findings, initially summarized by Zingales (1995) and subsequently more exhaustively by Frank and Goyal (2009)

- Consistent historic and cross-border evidence of leverage related to certain firm characteristics
- Depending on the interpretation of these firm characteristics, they play into virtually all the applied modeling papers
- Because the interpretations are different, can be used to support quite different theories

	Profitability	Market To Book	R&D	Tangibility	Selling Expenses	Risk	Size
Pecking order	−	+	+	−	+	+	−
Trade off	−	−	−	+	−	−	+
Frank and Goyal, 2009	−	−		+			+

Our purpose: Not going to propose a new mash-up

Look at contemporaneous correlation model (CCM) through the lens of structural causal modeling (SCM).

- An empirical framework to estimate causal effects from observational data

Our purpose: Not going to propose a new mash-up

Look at contemporaneous correlation model (CCM) through the lens of structural causal modeling (SCM).

- An empirical framework to estimate causal effects from observational data
- It forces you to write down an empirically testable model of how you think your variables are causally related

Our purpose: Not going to propose a new mash-up

Look at contemporaneous correlation model (CCM) through the lens of structural causal modeling (SCM).

- An empirical framework to estimate causal effects from observational data
- It forces you to write down an empirically testable model of how you think your variables are causally related
- With the empirical test you are not going to be able to reject (1) Many other possible SCMs (2) Reverse causation

Our purpose: Not going to propose a new mash-up

Look at contemporaneous correlation model (CCM) through the lens of structural causal modeling (SCM).

- An empirical framework to estimate causal effects from observational data
- It forces you to write down an empirically testable model of how you think your variables are causally related
- With the empirical test you are not going to be able to reject (1) Many other possible SCMs (2) Reverse causation
- But you are able to test if the data is consistent with your SCM

Revisit these relationships from a causal standpoint

Ex-ante identification of causal structures

Estimation of causal effects

e.g., Propensity Score, Regression Discontinuty, ...

Revisit these relationships from a causal standpoint

Ex-ante identification of causal structures

Estimation of causal effects

e.g., Propensity Score, Regression Discontinuty, ...

The importance of *Identification* is often underevalued and mostly performed ex-post based on quality of fit measures

Revisit these relationships from a causal standpoint

Ex-ante identification of causal structures

Estimation of causal effects

e.g., Propensity Score, Regression Discontinuty, ...

The importance of *Identification* is often underevalued and mostly performed ex-post based on quality of fit measures

Here we focus on causal structures identification using Structural Causal Modeling (SCM)

Revisit these relationships from a causal standpoint

Ex-ante identification of causal structures

Estimation of causal effects

e.g., Propensity Score, Regression Discontinuty, ...

The importance of *Identification* is often underevalued and mostly performed ex-post based on quality of fit measures

Here we focus on causal structures identification using Structural Causal Modeling (SCM)

- How are leverage and its determinants causally related?
- How do causal effects compare to the empirical associations?

Why is identification so important?

Y	X	S	R
y_1	x_1	s_1	r_1
y_2	x_2	s_2	r_2
y_3	x_3	s_3	r_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Why is identification so important?

$$y = \mathcal{F}(\dots)$$

Y	X	S	R
y_1	x_1	s_1	r_1
y_2	x_2	s_2	r_2
y_3	x_3	s_3	r_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Why is identification so important?

$$y = \mathcal{F}(\dots)$$

Y	X	S	R
y_1	x_1	s_1	r_1
y_2	x_2	s_2	r_2
y_3	x_3	s_3	r_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$

$\nearrow$ **Model A:** $y = \alpha x + \epsilon$

Why is identification so important?

$$y = \mathcal{F}(\dots)$$

Y	X	S	R
y_1	x_1	s_1	r_1
y_2	x_2	s_2	r_2
y_3	x_3	s_3	r_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Model A: $y = \alpha x + \epsilon$

Model B: $y = \alpha x + \beta r + \epsilon$

Why is identification so important?

$$y = \mathcal{F}(\dots)$$

Y	X	S	R
y_1	x_1	s_1	r_1
y_2	x_2	s_2	r_2
y_3	x_3	s_3	r_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Model A: $y = \alpha x + \epsilon$

Model B: $y = \alpha x + \beta r + \epsilon$

Model	α	$\sigma_{\bar{\alpha}}$	t-value	p-value	R^2	Residuals
Model A	**-0.00**	0.045	-0.009	0.502	0.001	Normal
Model B	**0.229**	0.037	6.099	0.000	0.346	Normal

The right model is the worst model

$$X \sim \mathcal{N}_x(0,1); S \sim \mathcal{N}_s(0,1);$$

$$Y \sim S + \mathcal{N}_y(0,1); R = -0.5X + 1.5S + \mathcal{N}_r(0,1)$$

Y	X	S	R
y_1	x_1	s_1	r_1
y_2	x_2	s_2	r_2
y_3	x_3	s_3	r_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Model A: $y = \alpha x + \epsilon$

Model B: $y = \alpha x + \beta r + \epsilon$

Model	α	$\sigma_{\bar{\alpha}}$	t-value	p-value	R^2	Residuals
Model A	**-0.00**	0.045	-0.009	0.502	0.001	Normal
Model B	0.229	0.037	6.099	0.000	0.346	Normal

Warning

A very simple dataset, two models, two coefficients but no statistical measure that helps you select the correct one

Connection to capital structure problems, Example

Size-Leverage relation

Trade-off theory, positive

Large, diversified firms have lower default risk and lower debt-related agency costs.

Pecking order, negative

Large firms have more retained earnings and have a lower cost of equity issuance.

Connection to capital structure problems, Example

Size-Leverage relation

Trade-off theory, positive

Large, diversified firms have lower default risk and lower debt-related agency costs.

Pecking order, negative

Large firms have more retained earnings and have a lower cost of equity issuance.

Generally positive, but negative after controlling for the choice of issuing rated debt (Faulkender, M. et al, 2005 and Hovakimian, A. et al, 2009,2011)

Connection to capital structure problems, Example

Size-Leverage relation

Trade-off theory, positive

Large, diversified firms have lower default risk and lower debt-related agency costs.

Pecking order, negative

Large firms have more retained earnings and have a lower cost of equity issuance.

Generally positive, but negative after controlling for the choice of issuing rated debt (Faulkender, M. et al, 2005 and Hovakimian, A. et al, 2009,2011)

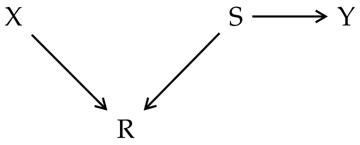
Warning (again) *two models, two coefficients, but no statistical measure that helps you select the correct one*

Let's go back to the numerical example

$$X \sim \mathcal{N}(0,1)$$
$$S \sim \mathcal{N}_s(0,1)$$
$$Y \sim S + \mathcal{N}_y(0,1)$$
$$R \sim -0.5X + 1.5S + \mathcal{N}_r(0,1)$$

Let's go back to the numerical example

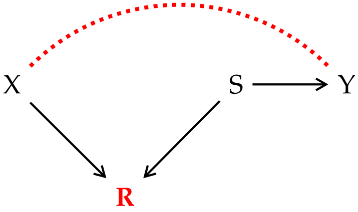
$$X \sim \mathcal{N}(0,1)$$
$$S \sim \mathcal{N}_s(0,1)$$
$$Y \sim S + \mathcal{N}_y(0,1)$$
$$R \sim -0.5X + 1.5S + \mathcal{N}_r(0,1)$$



Let's go back to the numerical example

$$X \sim \mathcal{N}(0,1)$$
$$S \sim \mathcal{N}_s(0,1)$$
$$Y \sim S + \mathcal{N}_y(0,1)$$
$$R \sim -0.5X + 1.5S + \mathcal{N}_r(0,1)$$

Collider bias



The problem of even a very small kitchen sink

a not-omitted variable bias

Estimating unbiased causal effects it's all about finding the right conditioning set. But how?

Y	X	Z	$\cdots$	$\cdots$	L
y_1	x_1	z_1			l_1
y_2	x_2	z_2			l_2
y_3	x_3	z_3			l_3
$\vdots$	$\vdots$	$\vdots$			$\vdots$

Estimating unbiased causal effects it's all about finding the right conditioning set. But how?

Y	X	Z	$\cdots$	$\cdots$	L
y_1	x_1	z_1			l_1
y_2	x_2	z_2			l_2
y_3	x_3	z_3			l_3
$\vdots$	$\vdots$	$\vdots$			$\vdots$

We need a story

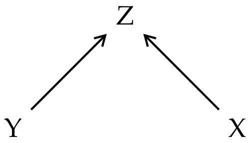
We don't know the model we only see the data

Y	X	Z
y_1	x_1	z_1
y_2	x_2	z_2
y_3	x_3	z_3
$\vdots$	$\vdots$	$\vdots$

We don't know the model we only see the data

Y	X	Z
y_1	x_1	z_1
y_2	x_2	z_2
y_3	x_3	z_3
$\vdots$	$\vdots$	$\vdots$

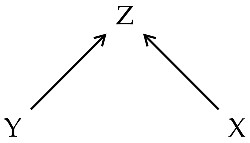
We can formulate hypotheses

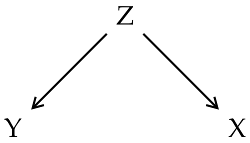


We don't know the model we only see the data

Y	X	Z
y_1	x_1	z_1
y_2	x_2	z_2
y_3	x_3	z_3
$\vdots$	$\vdots$	$\vdots$

We can formulate hypotheses

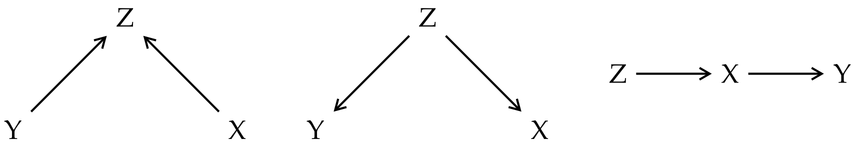




We don't know the model we only see the data

Y	X	Z
y_1	x_1	z_1
y_2	x_2	z_2
y_3	x_3	z_3
$\vdots$	$\vdots$	$\vdots$

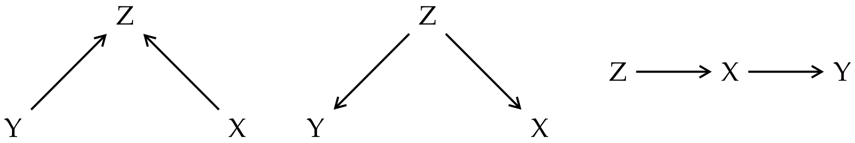
We can formulate hypotheses



We don't know the model we only see the data

Y	X	Z
y_1	x_1	z_1
y_2	x_2	z_2
y_3	x_3	z_3
$\vdots$	$\vdots$	$\vdots$

We can formulate hypotheses

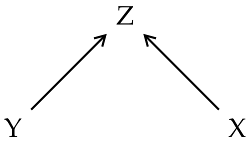


Z-control induces bias

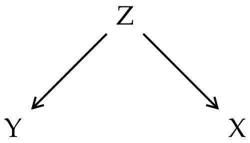
We don't know the model we only see the data

Y	X	Z
y_1	x_1	z_1
y_2	x_2	z_2
y_3	x_3	z_3
$\vdots$	$\vdots$	$\vdots$

We can formulate hypotheses



Z-control induces bias

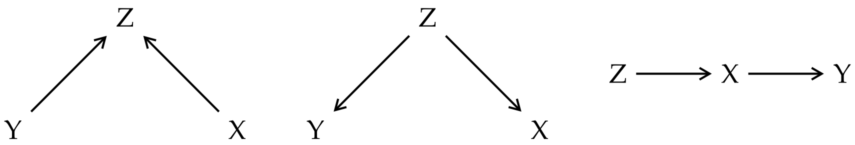


Z-control removes bias

We don't know the model we only see the data

Y	X	Z
y_1	x_1	z_1
y_2	x_2	z_2
y_3	x_3	z_3
$\vdots$	$\vdots$	$\vdots$

We can formulate hypotheses

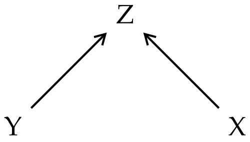


Z-control induces bias

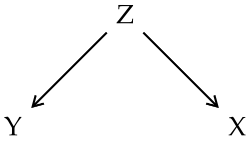
Z-control removes bias

Z-control irrelevant

How do we test our hypothesis?



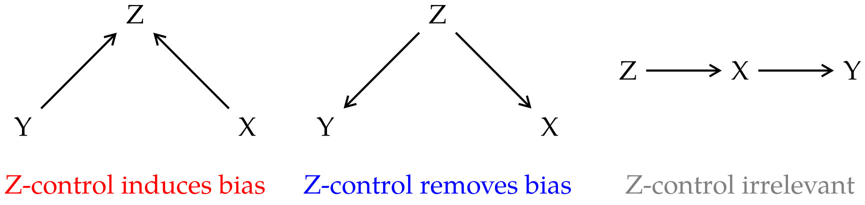
Z-control induces bias



Z-control removes bias

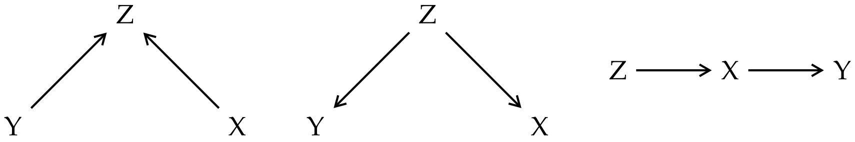
Z-control irrelevant

How do we test our hypothesis?



Each model or *causal graph* entails a set of conditional independences

How do we test our hypothesis?



Z-control induces bias

$X \perp\!\!\!\perp Y$

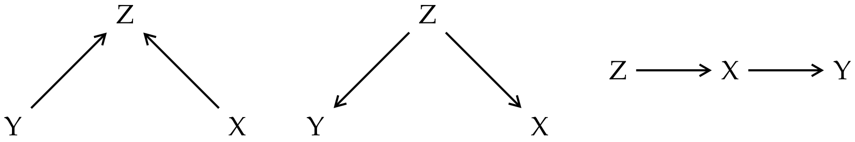
$X \not\perp\!\!\!\perp Y | Z$

Z-control removes bias

Z-control irrelevant

Each model or *causal graph* entails a set of conditional independences

How do we test our hypothesis?



Z-control induces bias

$X \perp\!\!\!\perp Y$

$X \not\perp\!\!\!\perp Y | Z$

Z-control removes bias

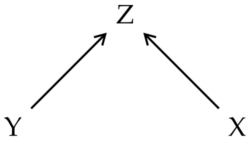
$X \not\perp\!\!\!\perp Y$

$X \perp\!\!\!\perp Y | Z$

Z-control irrelevant

Each model or *causal graph* entails a set of conditional independences

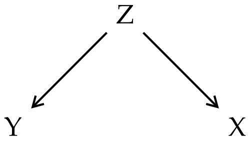
How do we test our hypothesis?



Z-control induces bias

$X \perp\!\!\!\perp Y$

$X \not\!\perp\!\!\!\perp Y | Z$



Z-control removes bias

$X \not\!\perp\!\!\!\perp Y$

$X \perp\!\!\!\perp Y | Z$

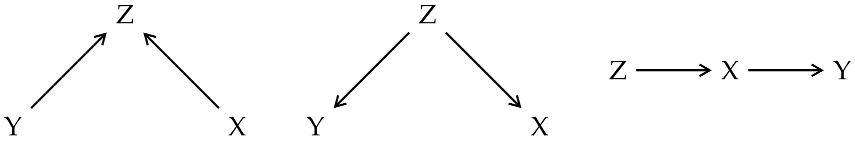
Z-control irrelevant

$X \not\!\perp\!\!\!\perp Y$

$X \not\!\perp\!\!\!\perp Y | Z$

Each model or *causal graph* entails a set of conditional independences

How do we test our hypothesis?

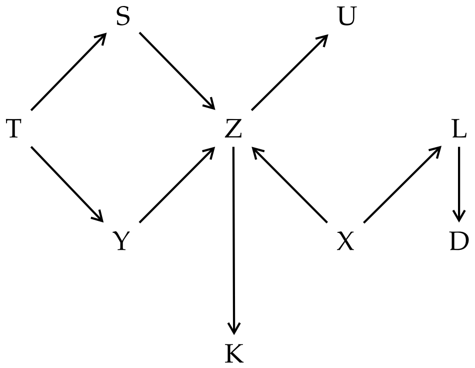


Z-control induces bias	Z-control removes bias	Z-control irrelevant
$X \perp\!\!\!\perp Y$	$X \not\perp\!\!\!\perp Y$	$X \not\perp\!\!\!\perp Y$
$X \not\perp\!\!\!\perp Y \mid Z$	$X \perp\!\!\!\perp Y \mid Z$	$X \not\perp\!\!\!\perp Y \mid Z$

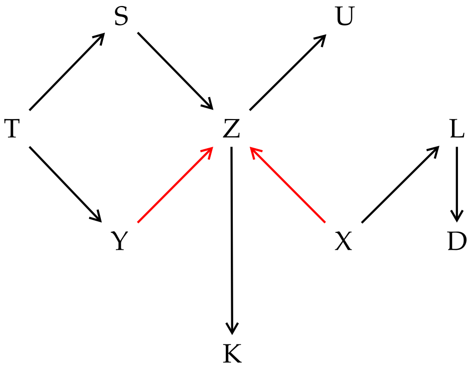
Each model or *causal graph* entails a set of conditional independences

Once we have the right graph we know exactly what to control for to estimate unbiased causal effects

Isn't the world more complex than these simple models?

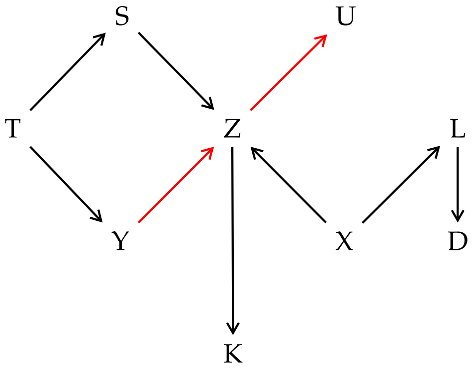


Isn't the world more complex than these simple models?



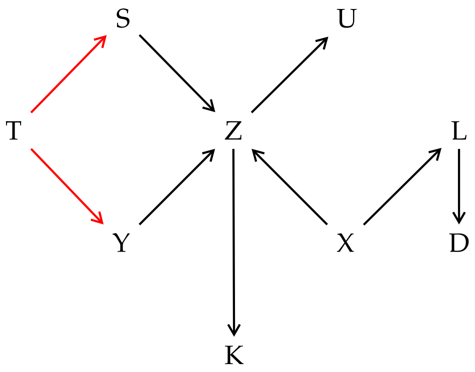
Collider: $X \perp\!\!\!\perp Y$, $X \not\perp\!\!\!\perp Y | Z$

Isn't the world more complex than these simple models?



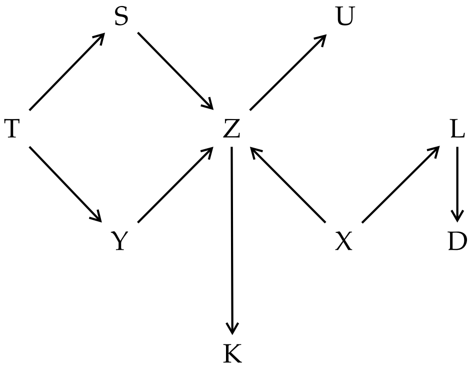
Chain: $Y \perp\!\!\!\perp U | Z$

Isn't the world more complex than these simple models?



Fork: $Y \perp\!\!\!\perp S | T$

Isn't the world more complex than these simple models?



The first step of a SCM is to determine whether the story we have in mind agrees with the data

d-separation

A path between two nodes X and Y is blocked by a set of nodes Z if and only if:

- The path contains a chain or a fork such that the middle node is in Z, or
- The path contain a collider such that the collision node is not in Z and no descendent of the collision node is in Z

if Z blocks every path between two nodes X and Y , then X and Y are d-separated conditional on Z. **Therefore they are conditionally independent given Z**

d-separation

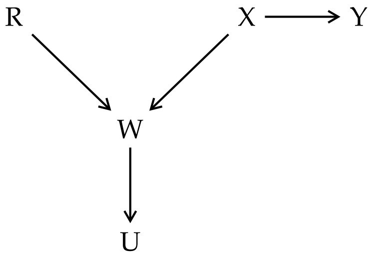
A path between two nodes X and Y is blocked by a set of nodes Z if and only if:

- The path contains a chain or a fork such that the middle node is in Z, or
- The path contain a collider such that the collision node is not in Z and no descendent of the collision node is in Z

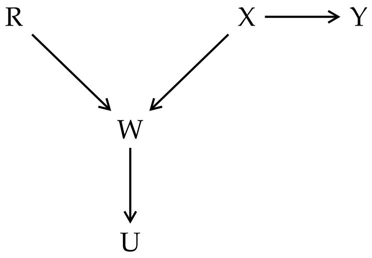
if Z blocks every path between two nodes X and Y , then X and Y are d-separated conditional on Z. **Therefore they are conditionally independent given Z**

The concept of d-separation allows us to determine the implied conditional independencies of a model

Example of d-separation

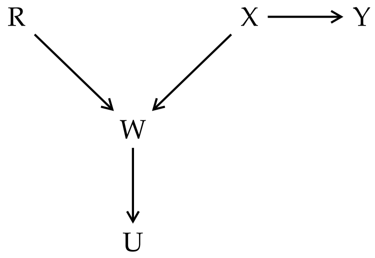


Example of d-separation



R and Y are d-separated using an empty set there is only one path between R and Y which is blocked by W. *They are unconditionally independent*

Example of d-separation



R and Y are d-separated using an empty set there is only one path between R and Y which is blocked by W. *They are unconditionally independent*

R and Y are d-connected if we condition on $\{W\}$ there is no chain or fork in the conditioning set (cond. 1 is not valid) and the only collider is in the conditioning set. *They are dependent conditioning on* $\{W\}$

Connection with regression

Structural causal models

- Formulate a hypothesis
- Test the hypothesis
- Use d-sep and backdoor to identify controls
- Estimate "unbiased" effects:
  1. Regression
  2. ...
  3. Do-operations (simulate an experiment)

Regression

- Formulate a hypothesis
- No real test you can do here
- Estimate the regression coefficients

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

In linear models with Gaussian distributions

$$x = \alpha z + \epsilon_x$$
$$y = \beta z + \epsilon_y$$

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

In linear models with Gaussian distributions

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$
$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

In linear models with Gaussian distributions

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$

$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Model (normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.000
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \vert Z$	0.763

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

In linear models with Gaussian distributions

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$

$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Model (normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.000
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \vert Z$	0.763
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.000
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \vert Z$	0.211

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

In linear models with Gaussian distributions

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$

$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Model (normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.000
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \| Z$	0.763
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.000
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \| Z$	0.211
Collider ($X \rightarrow Z \leftarrow Y$)	$X \perp\!\!\!\perp Y$	0.658
Collider ($X \rightarrow Z \leftarrow Y$)	$X \perp\!\!\!\perp Y \| Z$	0.000

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$

$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$

$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Model (non-normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \mid Z$	0

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$

$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Model (non-normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \mid Z$	0
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \mid Z$	0

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$x = \alpha z + \epsilon_x \qquad H_0 : \rho(\epsilon_x, \epsilon_y) = 0$$
$$y = \beta z + \epsilon_y \qquad H_1 : \rho(\epsilon_x, \epsilon_y) \neq 0$$

Model (non-normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \vert Z$	0
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \vert Z$	0
Collider ($X \rightarrow Z \leftarrow Y$)	$X \perp\!\!\!\perp Y$	0.48
Collider ($X \rightarrow Z \leftarrow Y$)	$X \perp\!\!\!\perp Y \vert Z$	0.15

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$X \perp\!\!\!\perp Y | Z \Leftrightarrow \Sigma_{\tilde{X}Y\cdot Z} = 0 \Leftrightarrow ||\Sigma_{\tilde{X}Y\cdot Z}||^2_{\text{HS}} = 0$$

$$\begin{aligned} H_0 &: ||\Sigma_{\tilde{X}Y\cdot Z}||^2_{\text{HS}} = 0 \\ H_1 &: ||\Sigma_{\tilde{X}Y\cdot Z}||^2_{\text{HS}} > 0 \end{aligned} \tag{1}$$

where $||\cdot||^2_{\text{HS}}$ is the Hilbert-Schmidt norm in Euclidean space.

Strobl, E. V., *et al.* (2019) *Journal of Causal Inference* → RCoT (Randomised conditional correlation test)

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$\begin{aligned} H_0 &: ||\Sigma_{\tilde{X}Y \cdot Z}||^2_{\text{HS}} = 0 \\ H_1 &: ||\Sigma_{\tilde{X}Y \cdot Z}||^2_{\text{HS}} > 0 \end{aligned} \tag{2}$$

Model (non-normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.03
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \mid Z$	0.45

Conditional independence tests

$$X \perp\!\!\!\perp Y | Z$$

For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$H_0 : ||\Sigma_{\tilde{X}Y \cdot Z}||^2_{\text{HS}} = 0$$
$$H_1 : ||\Sigma_{\tilde{X}Y \cdot Z}||^2_{\text{HS}} > 0 \quad (2)$$

Model (non-normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.03
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \| Z$	0.45
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.04
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y \| Z$	0.4

Conditional independence tests

$$X \perp\!\!\!\perp Y|Z$$

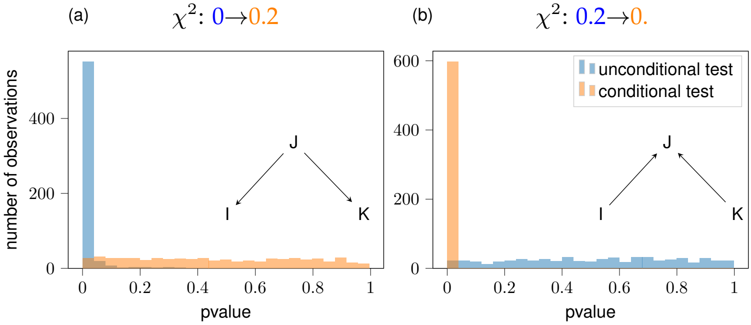
For non-Gaussian distributions zero partial correlation is neither necessary nor sufficient for conditional independence

$$H_0 : ||\Sigma_{\tilde{X}Y \cdot Z}||^2_{\text{HS}} = 0$$
$$H_1 : ||\Sigma_{\tilde{X}Y \cdot Z}||^2_{\text{HS}} > 0 \quad (2)$$

Model (non-normal noise)	Test	p-value
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.03
Fork ($X \leftarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y\|Z$	0.45
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y$	0.04
Chain ($X \rightarrow Z \rightarrow Y$)	$X \perp\!\!\!\perp Y\|Z$	0.4
Collider ($X \rightarrow Z \leftarrow Y$)	$X \perp\!\!\!\perp Y$	0.55
Collider ($X \rightarrow Z \leftarrow Y$)	$X \perp\!\!\!\perp Y\|Z$	0.05

Independence test on panel data

We run repeated tests for each year in the panel



A few observations

- SCM tells you which variables you need to control for to estimate causal effects

A few observations

- SCM tells you which variables you need to control for to estimate causal effects
- In econometrics, often a disconnect between the formal model and the empirical implementation

A few observations

- SCM tells you which variables you need to control for to estimate causal effects
- In econometrics, often a disconnect between the formal model and the empirical implementation
- SCM forces you to write down an explicit version of how the empirical variables you are using act on each other

A few observations

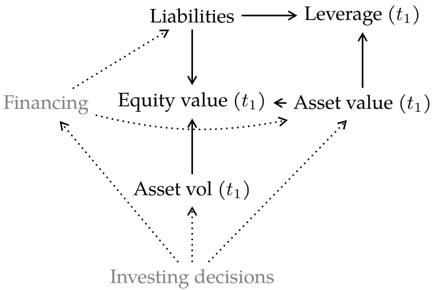
- SCM tells you which variables you need to control for to estimate causal effects
- In econometrics, often a disconnect between the formal model and the empirical implementation
- SCM forces you to write down an explicit version of how the empirical variables you are using act on each other
- The hypothesis are testable prior to estimation

Back to the problem: a causal model for leverage

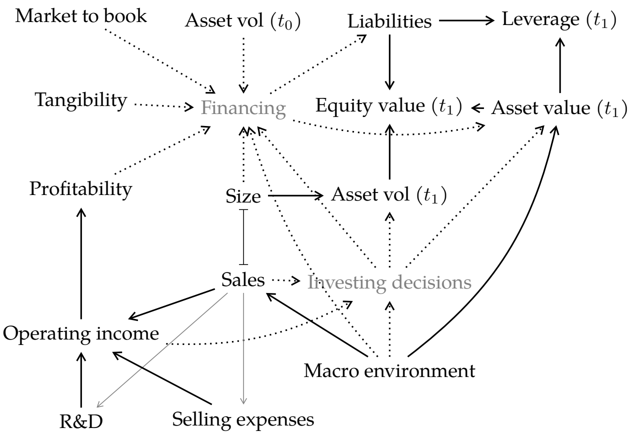
Financing

Investing decisions

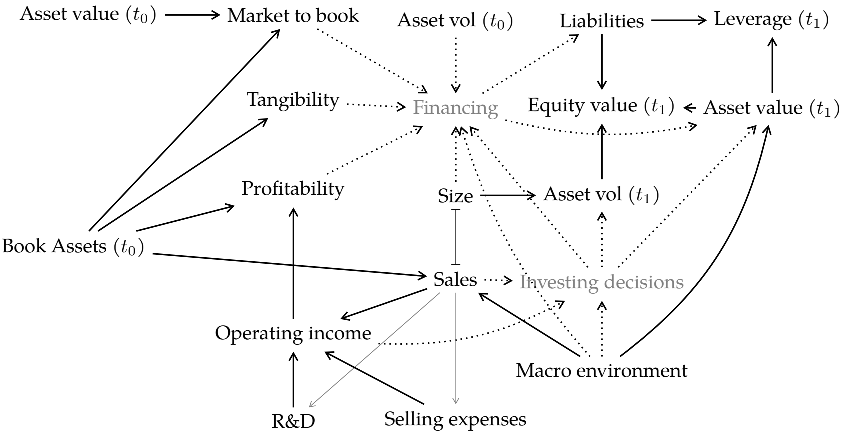
Back to the problem: a causal model for leverage

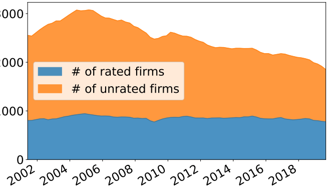


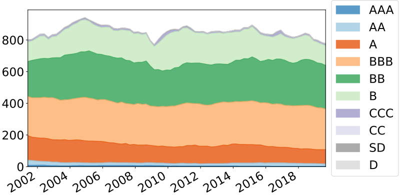
Back to the problem: a causal model for leverage



Back to the problem: a causal model for leverage

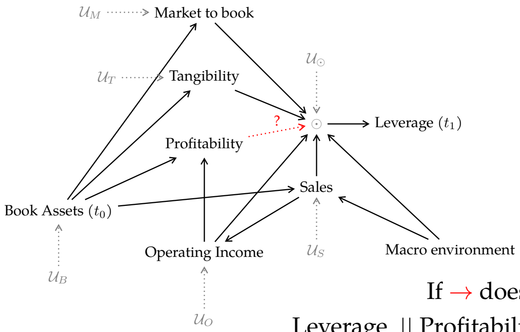






Variable	Non Rated	Rated
S&P 500 indicator	0.03	0.39
S&P 400 indicator	0.06	0.18
NYSE indicator	0.18	0.68
Probability rated	0.14	0.20
Market to book	1.83	1.69
Tangibility	0.26	0.37
R&D	0.05	0.02
Selling Expenses	0.31	0.17
Profitability	0.07	0.14
Size	4.50	7.68
Market Debt	0.12	0.26
Book Debt	0.26	0.37
Operating Risk	0.07	0.04
Market value of asset (log)	7.2	9.89
Volatility of asset	0.33	0.18
Observations	120837	62473

Testing conditional independencies

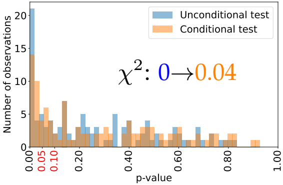


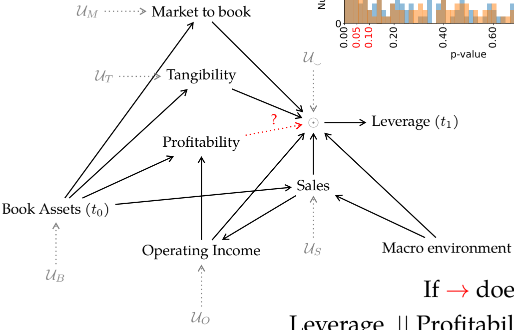
If $\rightarrow$ does not exists:

Leverage $\perp\!\!\!\perp$ Profitability | Book asset, Op. Inc.

Leverage $\perp\!\!\!\perp$ Profitability | Op. Inc., MTB, Sales, Tang.

Testing conditional independencies





If → does not exists:

Leverage $\perp\!\!\!\perp$ Profitability | Book asset, Op. Inc.

Leverage $\perp\!\!\!\perp$ Profitability | Op. Inc., MTB, Sales, Tang.

Testing conditional independencies

Hypothesis	χ^2
Profitability $\perp\!\!\!\perp$ Leverage|Sales,Op. Inc.,Market to book,Tangibility,Macro environment	0.0→0.04
Market to book $\perp\!\!\!\perp$ Leverage|Book assets$_{t_0}$	0.0→0.0
Market to book $\perp\!\!\!\perp$ Leverage|Sales,Op. Inc.,Profitability,Tangibility,Macro environment	0.0→0.0
Selling Expenses $\perp\!\!\!\perp$ Leverage|Sales	0.0→0.0
Selling Expenses $\perp\!\!\!\perp$ Leverage|Sales,Op. Inc.	0.0→0.0
R&D $\perp\!\!\!\perp$ Leverage|Sales	0.0→0.26
R&D $\perp\!\!\!\perp$ Leverage|Sales,Op. Inc.	0.0→0.26
Tangibility $\perp\!\!\!\perp$ Leverage|Book assets$_{t_0}$	0.0→0.0
Tangibility $\perp\!\!\!\perp$ Leverage|Sales,Op. Inc.,Profitability,Market to book,Macro environment	0.0→0.0
OpRisk $\perp\!\!\!\perp$ Leverage|Op. Inc.	0.41→0.61
Sales $\perp\!\!\!\perp$ Leverage|Op. Inc.,Profitability,Market to book,Tangibility,Macro environment	0.0→0.0
Sales $\perp\!\!\!\perp$ Leverage|Op. Inc.,Book assets$_{t_0}$,Macro environment	0.0→0.0
Risk $\perp\!\!\!\perp$ Leverage|Asset value$_{t_0}$,Book assets$_{t_0}$	0.0→0.0

Comparing estimated effects

	Statistical model (without rating control)	Causal model
Market to book	-0.240	-0.553
Tangibility	0.221	0.300
Selling expenses	-0.085	-0.576
Profitability	-0.353	-0.353
Size	0.139	0.581
Operating risk	0.013	NaN
Risk	NaN	-0.809

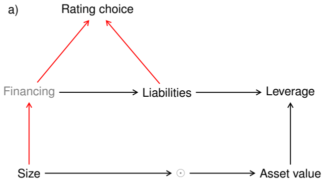
Comparing estimated effects

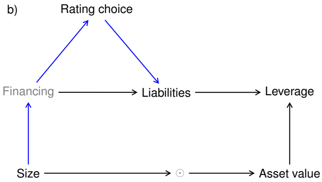
	Statistical model (without rating control)	Causal model
Market to book	-0.240	-0.553
Tangibility	0.221	0.300
Selling expenses	-0.085	-0.576
Profitability	-0.353	-0.353
Size	0.139	0.581
Operating risk	0.013	NaN
Risk	NaN	-0.809
	Mixed effects	Total effects

What is the role of the rating choice?

	Statistical model (with rating control)	Statistical model (without rating control)	Causal model
Market to book	-0.184	-0.240	-0.553
Tangibility	0.191	0.221	0.300
Selling expenses	-0.080	-0.085	-0.576
Profitability	-0.305	-0.353	-0.353
Size	-0.076	0.139	0.581
Operating risk	-0.002	0.013	NaN
Risk	NaN	NaN	-0.809

What is the role of the rating choice?

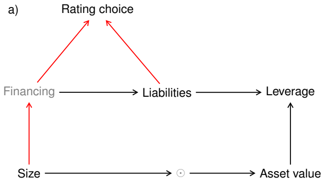


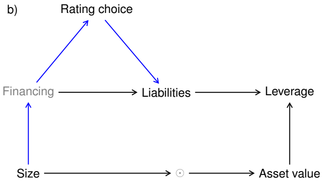


Collider

Mediator

What is the role of the rating choice?





Collider

Mediator

Hard to imagine a model with the rating choice as part of the backdoor path

Take home messages

- Capital structure theories

Take home messages

- Capital structure theories

  1. We derived and validated a SCM for leverage

Take home messages

- Capital structure theories
  1. We derived and validated a SCM for leverage
  2. Results provide support for the causal role of variables that measure the potential for information asymmetry

Take home messages

- Capital structure theories

  1. We derived and validated a SCM for leverage
  2. Results provide support for the causal role of variables that measure the potential for information asymmetry
  3. An explanation of the size-leverage controversy

Take home messages

- Capital structure theories
  1. We derived and validated a SCM for leverage
  2. Results provide support for the causal role of variables that measure the potential for information asymmetry
  3. An explanation of the size-leverage controversy
- Broader implications

Take home messages

- Capital structure theories
  1. We derived and validated a SCM for leverage
  2. Results provide support for the causal role of variables that measure the potential for information asymmetry
  3. An explanation of the size-leverage controversy
- Broader implications
  1. Identification of causal structures and ex-ante model validation

Take home messages

- Capital structure theories
  1. We derived and validated a SCM for leverage
  2. Results provide support for the causal role of variables that measure the potential for information asymmetry
  3. An explanation of the size-leverage controversy
- Broader implications
  1. Identification of causal structures and ex-ante model validation
  2. Multiple specifications to estimate comparable effects

Take home messages

- Capital structure theories
  1. We derived and validated a SCM for leverage
  2. Results provide support for the causal role of variables that measure the potential for information asymmetry
  3. An explanation of the size-leverage controversy
- Broader implications
  1. Identification of causal structures and ex-ante model validation
  2. Multiple specifications to estimate comparable effects
  3. Risk of introducing uncontrolled control variables (Unomitted variable problem)

Limitations

- Our approach does not solve omitted variables problems.

Limitations

- Our approach does not solve omitted variables problems.
- Cannot deal with cycles in the model.

Limitations

- Our approach does not solve omitted variables problems.
- Cannot deal with cycles in the model.
- There are identifiability issues since a given set of conditional independences can be shared by multiple graphs.

Limitations

- Our approach does not solve omitted variables problems.
- Cannot deal with cycles in the model.
- There are identifiability issues since a given set of conditional independences can be shared by multiple graphs.
- The estimation step is rudimentary

Limitations

- Our approach does not solve omitted variables problems.
- Cannot deal with cycles in the model.
- There are identifiability issues since a given set of conditional independences can be shared by multiple graphs.
- The estimation step is rudimentary
- The relevance of the estimated causal links hinges upon the economic interpretations of the empirical variables