Regression Discontinuity

Prof. Mike Aguilar

https:/www.linkedin.com/in/mike-aguilar-econ

Outline

1. Case Study
2. Incorrectly Estimating TE
3. Sharp RD
4. Fuzzy RD

Regression Discontinuity - Summary

- Natural Experiment

Regression Discontinuity - Summary

- Natural Experiment
- If the treatment is assigned according to a rule, then Regression Discontinuity (RD) might be able to reveal the ATE

Regression Discontinuity - Summary

- Natural Experiment
- If the treatment is assigned according to a rule, then Regression Discontinuity (RD) might be able to reveal the ATE
- Sharp RD
  - Treatment is a deterministic, discontinuous function of "running" variable
  - Regress outcome on treatment and running variable

Regression Discontinuity - Summary

- Natural Experiment
- If the treatment is assigned according to a rule, then Regression Discontinuity (RD) might be able to reveal the ATE
- Sharp RD
  - Treatment is a deterministic, discontinuous function of "running" variable
  - Regress outcome on treatment and running variable
- Fuzzy RD
  - Treatment is assigned by more than just the running variable
  - Using a threshold dummy as an instrument for the treatment dummy generates an unbiased IV estimate of the ATE

Housekeeping

- Throughout the slide deck you will see "Q", which indicates a question to you, the reader
- You will also see "A", which indicates the associated answer
- It is generally most efficient to learn this material through active participation. Whenever you encounter a "Q", be sure to try and develop your answer before turning the page to the provided "A" answer
- I recommend reading through these slides before engaging in associated coding exercises
- This slide deck is part of a broader curriculum in causal estimation and inference
- Suggested pre-requisite: instrument variables

Outline

1. Case Study
2. Incorrectly Estimating TE
3. Sharp RD
4. Fuzzy RD

Big Advisor Advantage

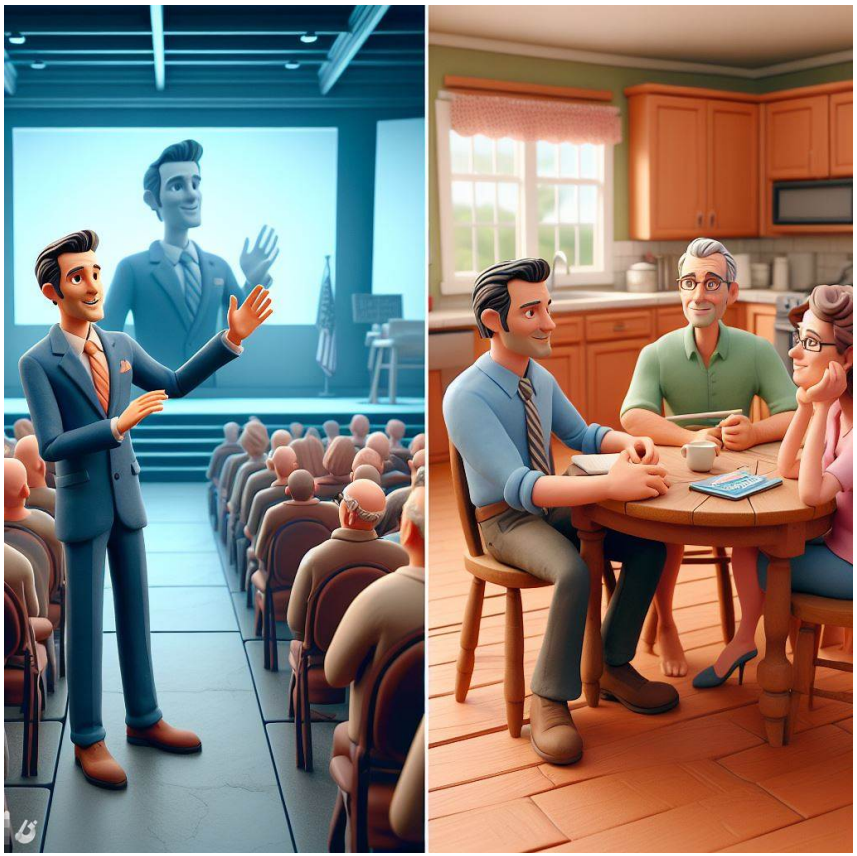


Figure: Image created with Bing Image Generator

Big Advisor Advantage

- Financial advisors (aka RIAs) are heavily regulated.

Big Advisor Advantage

- Financial advisors (aka RIAs) are heavily regulated.
- The regulating body is determined by the size of the assets the advisor manages, as measured by AUM (Assets Under Management).

Big Advisor Advantage

- Financial advisors (aka RIAs) are heavily regulated.
- The regulating body is determined by the size of the assets the advisor manages, as measured by AUM (Assets Under Management).
- Advisors with AUM $< \$100 Mln$ are regulated by State entities, whereas those with AUM $\geq \$100 Mln$ are regulated by Federal entities (i.e. the SEC).

Big Advisor Advantage

- Financial advisors (aka RIAs) are heavily regulated.
- The regulating body is determined by the size of the assets the advisor manages, as measured by AUM (Assets Under Management).
- Advisors with AUM $< \$100 Mln$ are regulated by State entities, whereas those with AUM $\geq \$100 Mln$ are regulated by Federal entities (i.e. the SEC).
- SEC oversight is more burdensome.

Big Advisor Advantage

- Financial advisors (aka RIAs) are heavily regulated.
- The regulating body is determined by the size of the assets the advisor manages, as measured by AUM (Assets Under Management).
- Advisors with AUM $< \$100 Mln$ are regulated by State entities, whereas those with AUM $\geq \$100 Mln$ are regulated by Federal entities (i.e. the SEC).
- SEC oversight is more burdensome.
- RIA revenues are often tied to AUM, with revenue collected measured as a percentage of assets managed; i.e. more AUM, more revenue

The Tradeoff

- Advisors face a tradeoff

The Tradeoff

- Advisors face a tradeoff
  - Grow above $\$100 Mln$ in pursuit of greater revenues and succumb to onerous SEC oversight, or

The Tradeoff

- Advisors face a tradeoff
  - Grow above $\$100 Mln$ in pursuit of greater revenues and succumb to onerous SEC oversight, or
  - Stay small with less regulation

The Tradeoff

- Advisors face a tradeoff
  - Grow above \$100*Mln* in pursuit of greater revenues and succumb to onerous SEC oversight, or
  - Stay small with less regulation
- Case for growing big: It is believed by many RIAs that clients respond positively to having the SEC "stamp of approval" that comes with the large size.

The Tradeoff

- Advisors face a tradeoff
  - Grow above $\$100 Mln$ in pursuit of greater revenues and succumb to onerous SEC oversight, or
  - Stay small with less regulation
- Case for growing big: It is believed by many RIAs that clients respond positively to having the SEC "stamp of approval" that comes with the large size.
- Case for staying small: Others believe that clients prefer a smaller, more personal firm.

The Tradeoff

- Advisors face a tradeoff
  - Grow above $\$100 Mln$ in pursuit of greater revenues and succumb to onerous SEC oversight, or
  - Stay small with less regulation
- Case for growing big: It is believed by many RIAs that clients respond positively to having the SEC "stamp of approval" that comes with the large size.
- Case for staying small: Others believe that clients prefer a smaller, more personal firm.
- The question for the RIA is how much the SEC "stamp of approval" will impact revenues.

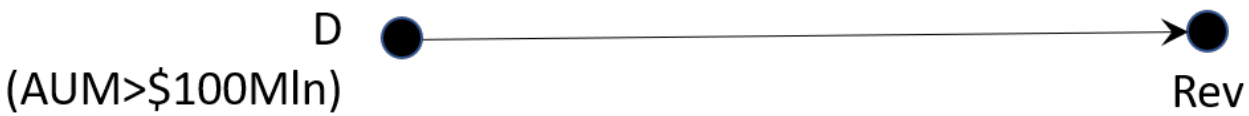
The Data

- Several thousand RIAs throughout the U.S.
- A single RIA is the observation unit (i)
- Data is measured for the year 2018
- Revenues are measured in $\$Mln$
- AUM is measured in $\$Mln$
- For each (i) we also create/have a dummy variable D $= 1$ if AUM $\geq$ \$100Mln and 0 if $<$ \$100Mln

Outcome? Treatment? Treated/Control? Treatment Effect? DAG?

Outcome? Treatment? Treated/Control? Treatment Effect? DAG?

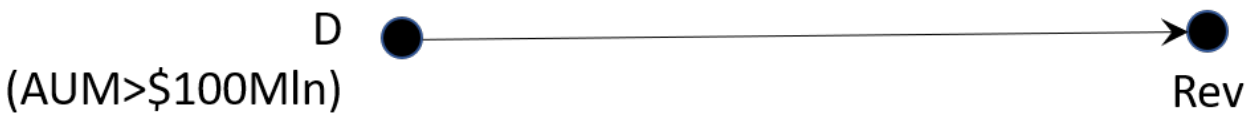
A:



- Outcome:

Outcome? Treatment? Treated/Control? Treatment Effect? DAG?

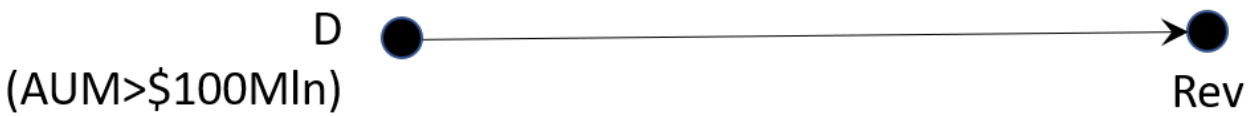
A:



- Outcome: Revenue
- Treatment:

Outcome? Treatment? Treated/Control? Treatment Effect? DAG?

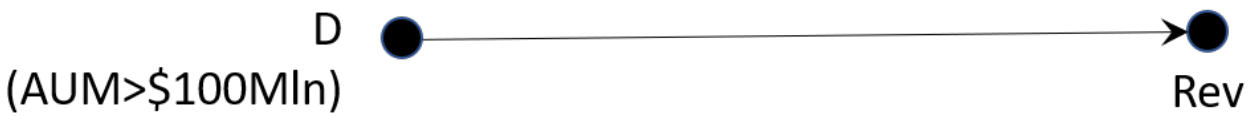
A:



- Outcome: Revenue
- Treatment:AUM$>$ \$100*Mln* (SEC oversight)
- Treated Group:

Outcome? Treatment? Treated/Control? Treatment Effect? DAG?

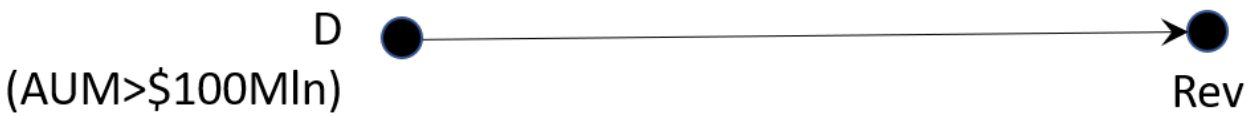
A:



- Outcome: Revenue
- Treatment:AUM$>$ \$100*Mln* (SEC oversight)
- Treated Group: RIAs with AUM$\geq$ \$100*Mln*
- Control Group:

Outcome? Treatment? Treated/Control? Treatment Effect? DAG?

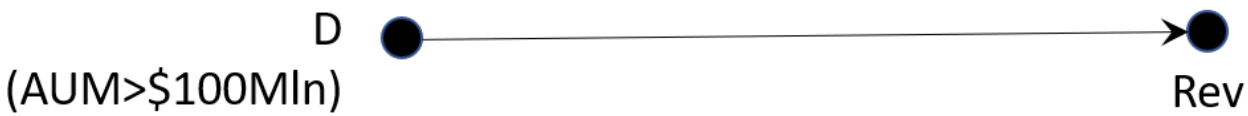
A:



- Outcome: Revenue
- Treatment:AUM$>$ \$100*Mln* (SEC oversight)
- Treated Group: RIAs with AUM$\geq$ \$100*Mln*
- Control Group: RIAs with AUM$<$ \$100*Mln*
- Treatment Effect:

Outcome? Treatment? Treated/Control? Treatment Effect? DAG?

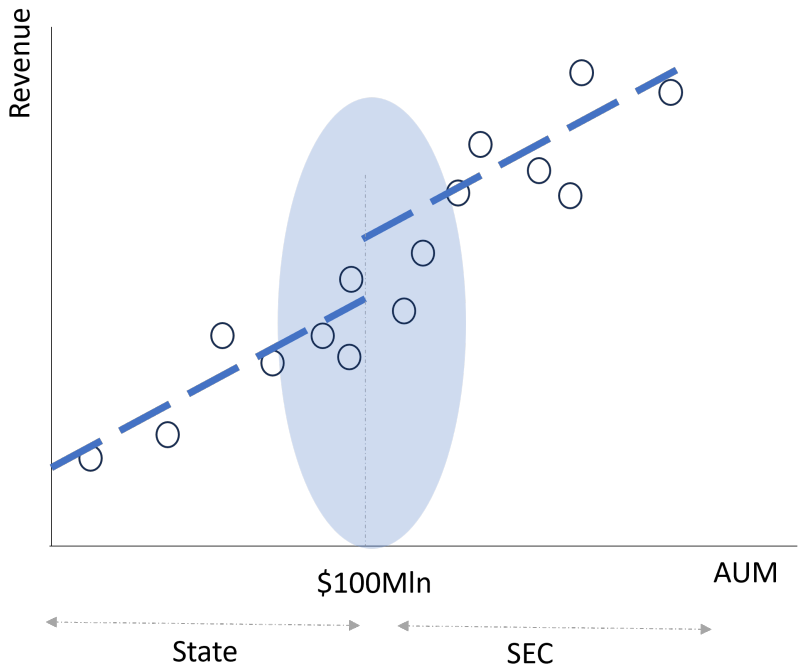
A:

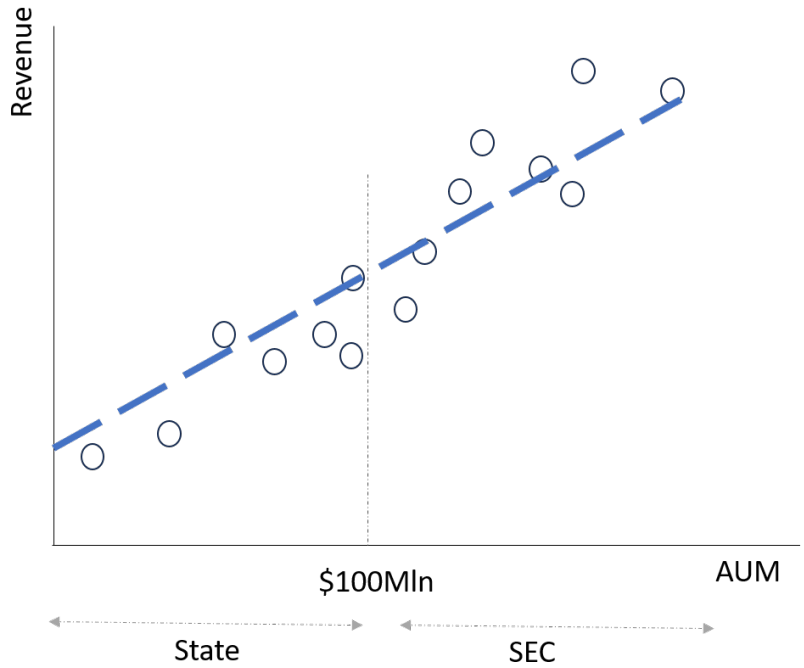


- Outcome: Revenue
- Treatment:AUM$>$ \$100*Mln* (SEC oversight)
- Treated Group: RIAs with AUM$\geq$ \$100*Mln*
- Control Group: RIAs with AUM$<$ \$100*Mln*
- Treatment Effect: impact of AUM $>$ \$100*Mln* upon Revenue

EDA

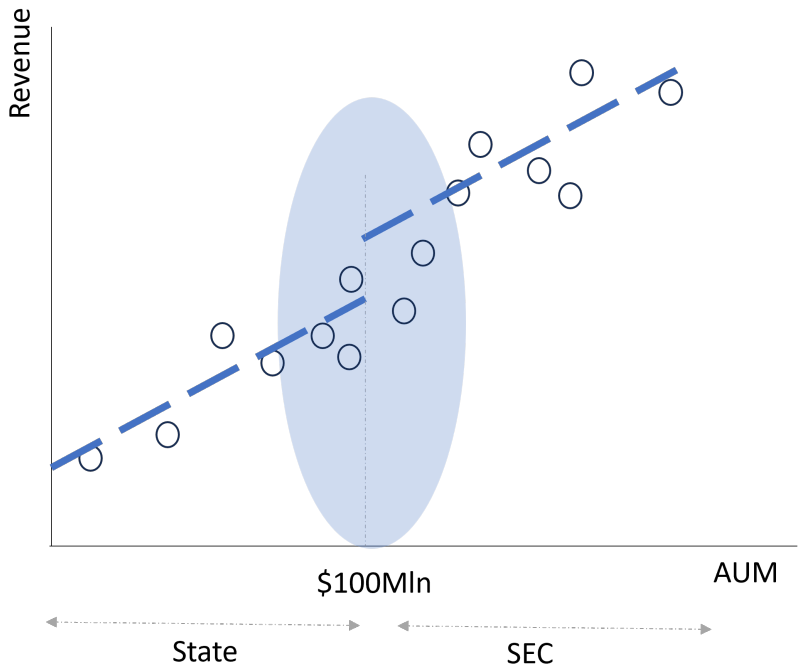
It's difficult to tell if there is a break at $\$100 Mln$ or just random noise

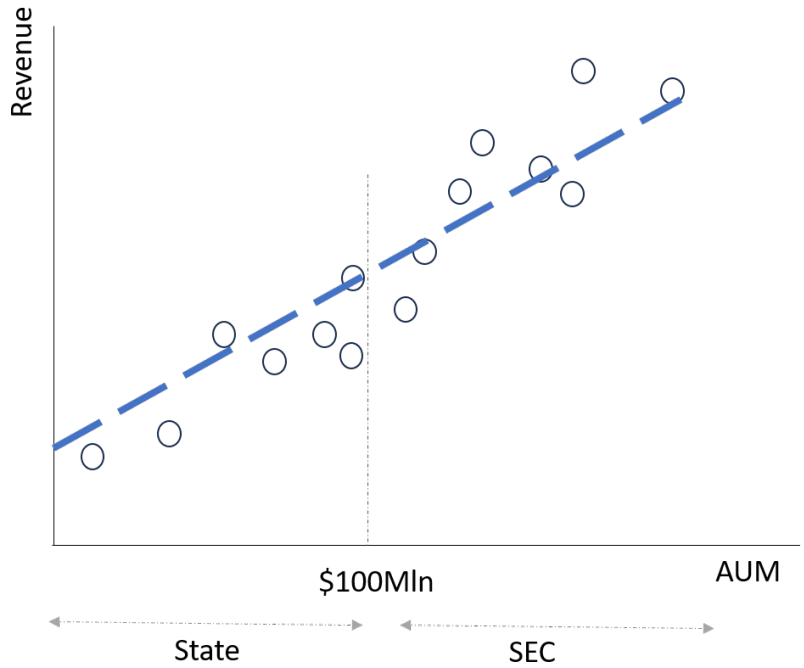




EDA

It's difficult to tell if there is a break at $\$100 Mln$ or just random noise

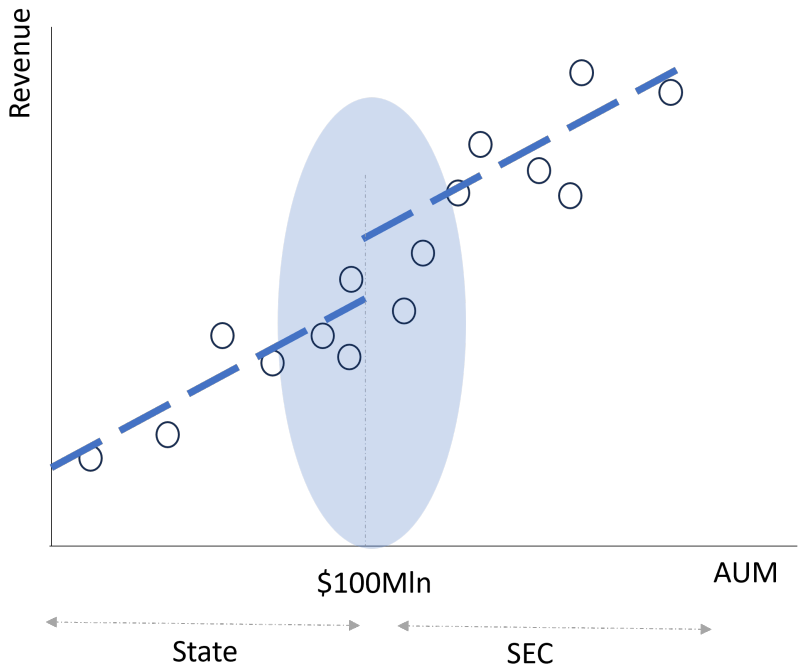




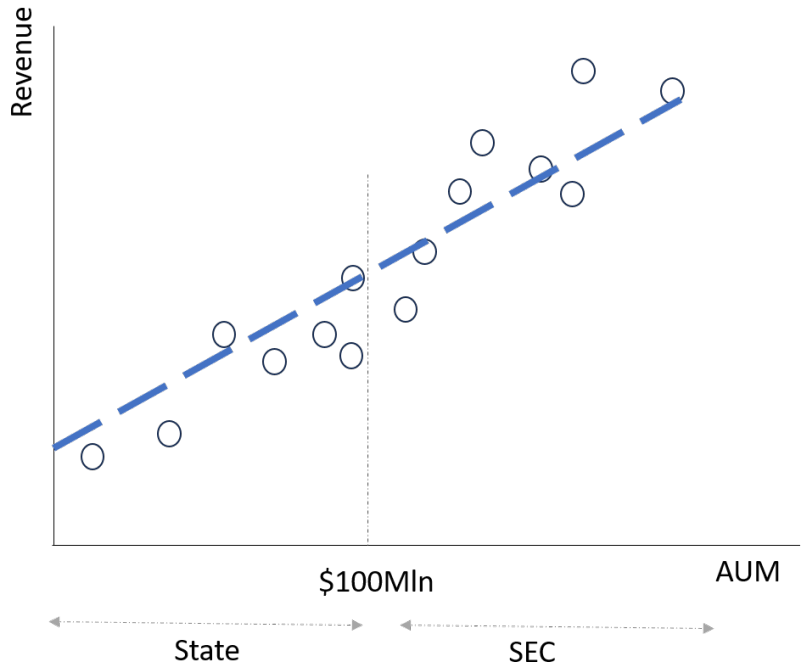
Do we have an "SEC Effect"? Or...

EDA

It's difficult to tell if there is a break at $\$100 Mln$ or just random noise



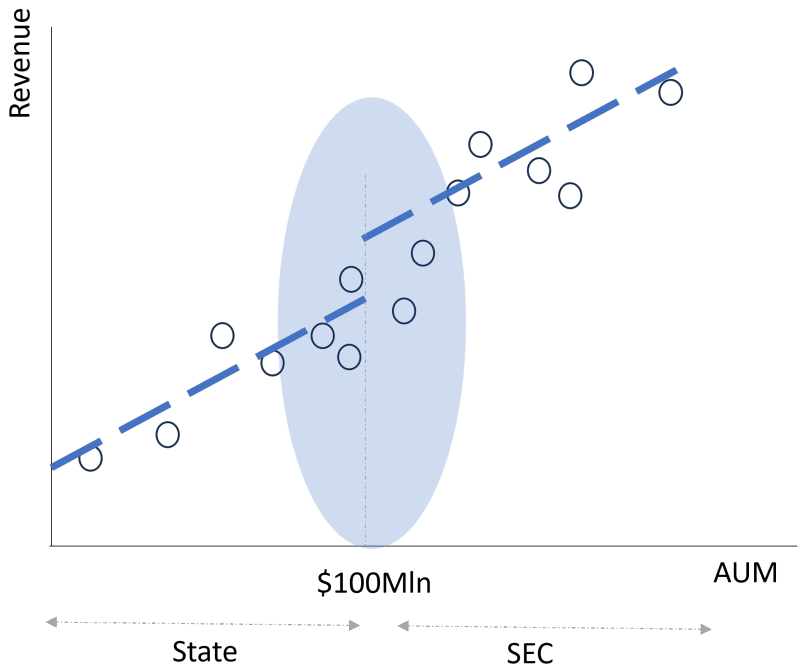
Do we have an "SEC Effect"? Or...



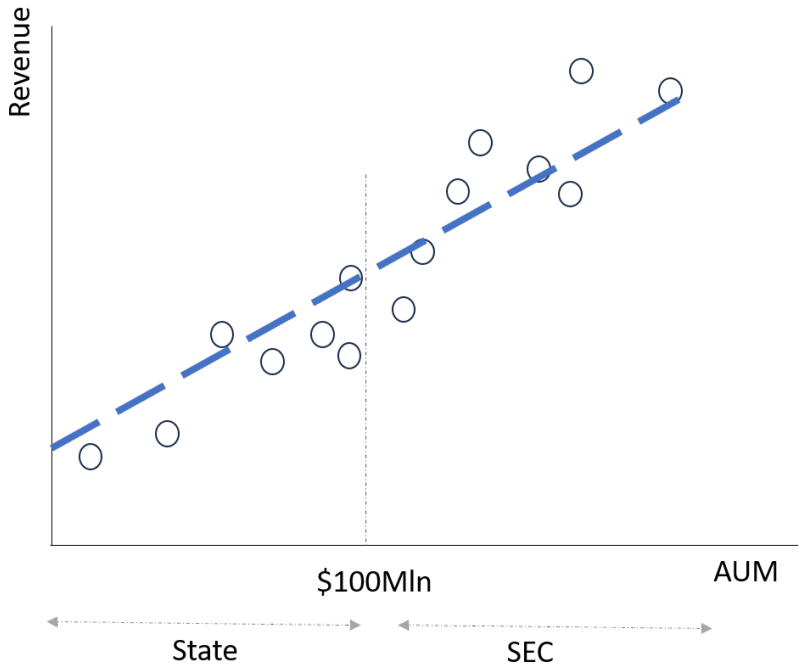
Do we have a noisy "Size Effect"?

EDA

It's difficult to tell if there is a break at $100*Mln* or just random noise



Do we have an "SEC Effect"? Or...

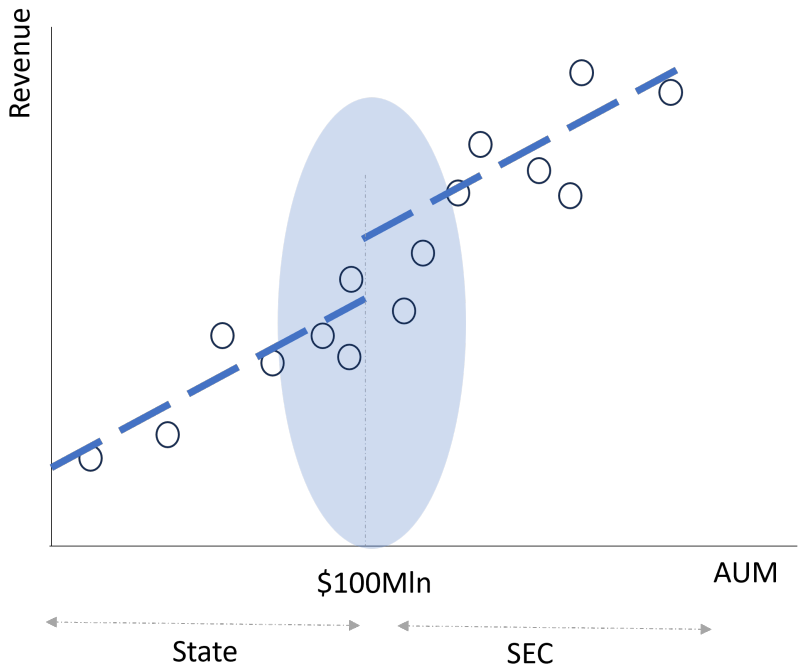


Do we have a noisy "Size Effect"?

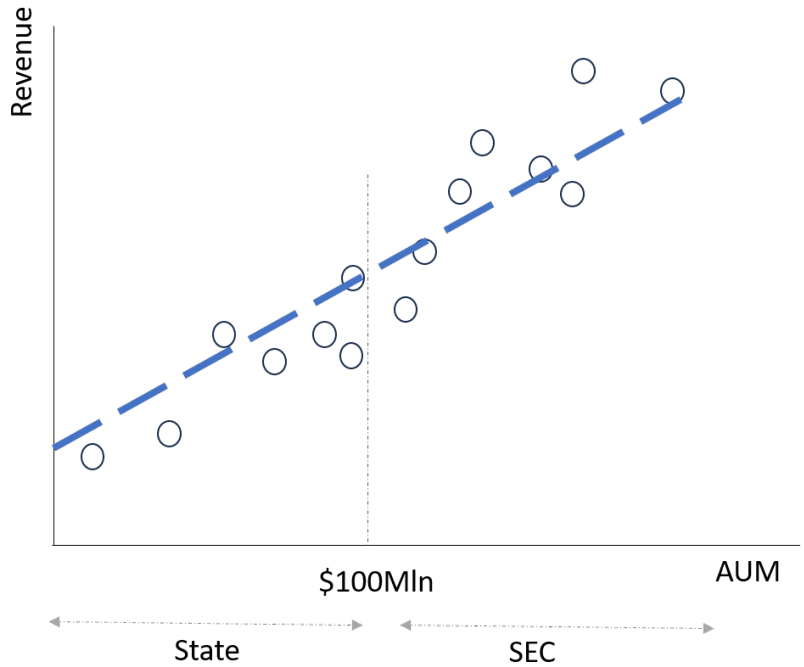
Q: What if the Client phrases the question as "Should we grow above $100Mln?" What is our setup missing?

EDA

It's difficult to tell if there is a break at $100*Mln* or just random noise



Do we have an "SEC Effect"? Or...



Do we have a noisy "Size Effect"?

Q: What if the Client phrases the question as "Should we grow above $100Mln?" What is our setup missing?
A: Costs, and Profit

Outline

1 Case Study

2 Incorrectly Estimating TE

3 Sharp RD

4 Fuzzy RD

TE via DiM

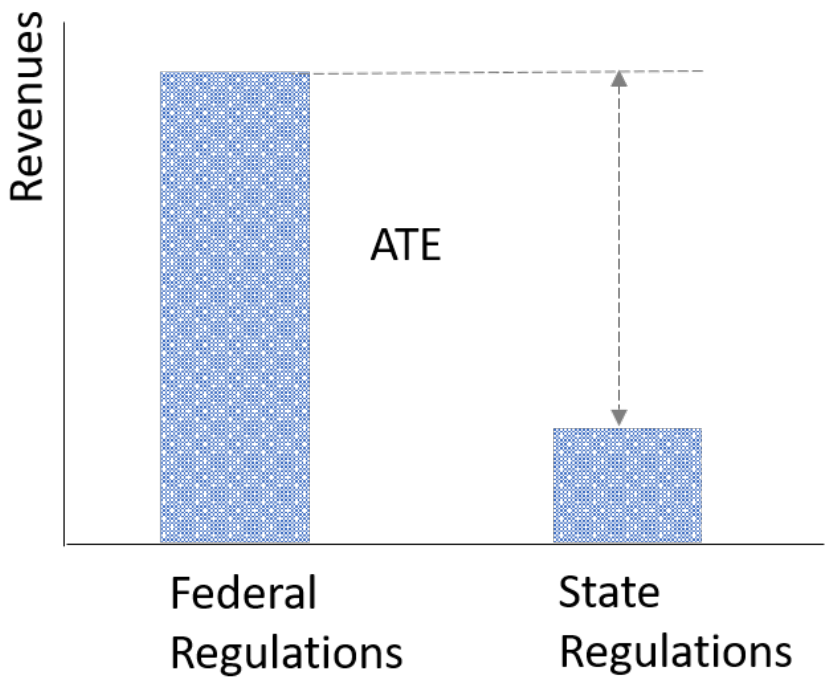
Q: How structure a simple DiM estimator for TE?

TE via DiM

Q: How structure a simple DiM estimator for TE?

A:

$$\widehat{ATE} = (\overline{Rev}|D = 1) - (\overline{Rev}|D = 0)$$



TE via OLS

Q: How replicate the DiM estimator via OLS?

TE via OLS

Q: How replicate the DiM estimator via OLS?

A:

$$Rev_i = \alpha + \beta D_i + e_i$$

From this regression we can estimate

TE via OLS

Q: How replicate the DiM estimator via OLS?

A:

$$Rev_i = \alpha + \beta D_i + e_i$$

From this regression we can estimate

$$\begin{aligned}
\widehat{ATE}^{OLS} &= \hat{\beta} \\
&= (\overline{Rev}|D = 1) - (\overline{Rev}|D = 0) \\
&= \widehat{ATE}^{DiM}
\end{aligned}$$

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

A:

$$(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$$

$$= (\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1) +$$
$$(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$$

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

A:

$$(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$$

$$= (\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1) +$$
$$(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$$

- $\widehat{ATT}$:

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

A:

$$(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$$

$$= (\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1) + (\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$$

- $\widehat{ATT}$: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

A:

$$(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$$

$$= (\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1) +$$
$$(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$$

- $\widehat{ATT}$: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$
- ATT:

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

A:

$$(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$$

$$= (\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1) +$$
$$(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$$

- $\widehat{ATT}$: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$
- ATT: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1)$

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

A:

$$(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$$

$$= (\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1) +$$
$$(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$$

- $\widehat{ATT}$: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$
- ATT: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1)$
- Selection Bias:

Isolating the Selection Bias

Q: Can you isolate the Selection Bias of our DiM estimator?

A:

$$(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$$

$$= (\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1) +$$
$$(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$$

- $\widehat{ATT}$: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=0)$
- ATT: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1)$
- Selection Bias: $(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$

Explaining the Selection Bias

Q: Can you explain the ATT and selection bias in our case?

Explaining the Selection Bias

Q: Can you explain the ATT and selection bias in our case?

A:

ATT: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1)$

Explaining the Selection Bias

Q: Can you explain the ATT and selection bias in our case?

A:

ATT: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1)$

- How much more/less BIG firms would have earned by being regulated by SEC rather than State authorities

Explaining the Selection Bias

Q: Can you explain the ATT and selection bias in our case?

A:

ATT: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1)$

- How much more/less BIG firms would have earned by being regulated by SEC rather than State authorities

Selection Bias: $(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$

Explaining the Selection Bias

Q: Can you explain the ATT and selection bias in our case?

A:

ATT: $(\overline{Rev}_1|D=1) - (\overline{Rev}_0|D=1)$

- How much more/less BIG firms would have earned by being regulated by SEC rather than State authorities

Selection Bias: $(\overline{Rev}_0|D=1) - (\overline{Rev}_0|D=0)$

- How much more/less BIG firms would have earned than small if they were under State regulations

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

We can compute the expected outcome in treated and untreated states

$$E[Rev_i|D_i = 1] =$$
$$E[Rev_i|D_i = 0] =$$

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

We can compute the expected outcome in treated and untreated states

$$E[Rev_i|D_i = 1] = \alpha + \beta + E[e_i|D_i = 1]$$
$$E[Rev_i|D_i = 0] =$$

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

We can compute the expected outcome in treated and untreated states

$$E[Rev_i|D_i = 1] = \alpha + \beta + E[e_i|D_i = 1]$$
$$E[Rev_i|D_i = 0] = \alpha + E[e_i|D_i = 0]$$

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

We can compute the expected outcome in treated and untreated states

$$E[Rev_i|D_i = 1] = \alpha + \beta + E[e_i|D_i = 1]$$
$$E[Rev_i|D_i = 0] = \alpha + E[e_i|D_i = 0]$$

Then the ATE is

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

We can compute the expected outcome in treated and untreated states

$$E[Rev_i|D_i = 1] = \alpha + \beta + E[e_i|D_i = 1]$$
$$E[Rev_i|D_i = 0] = \alpha + E[e_i|D_i = 0]$$

Then the ATE is

$$E[Rev_i|D_i = 1] - E[Rev_i|D_i = 0] =$$

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

We can compute the expected outcome in treated and untreated states

$$E[Rev_i|D_i = 1] = \alpha + \beta + E[e_i|D_i = 1]$$
$$E[Rev_i|D_i = 0] = \alpha + E[e_i|D_i = 0]$$

Then the ATE is

$$E[Rev_i|D_i = 1] - E[Rev_i|D_i = 0] = \beta$$
$$+$$

Isolating the OVB

Q: Can you isolate the OVB of our OLS estimator?

A: Given the model

$$Rev_i = \alpha + \beta D_i + e_i$$

We can compute the expected outcome in treated and untreated states

$$\begin{aligned} E[Rev_i|D_i = 1] &= \alpha + \beta + E[e_i|D_i = 1] \\ E[Rev_i|D_i = 0] &= \alpha + E[e_i|D_i = 0] \end{aligned}$$

Then the ATE is

$$\begin{aligned} E[Rev_i|D_i = 1] - E[Rev_i|D_i = 0] &= \beta \\ &+ (E[e_i|D_i = 1] - E[e_i|D_i = 0]) \end{aligned}$$

Explaining the OVB

Q: Can you explain the OVB in our case? Can you guess a sign?

Explaining the OVB

Q: Can you explain the OVB in our case? Can you guess a sign?

A: OVB: $(E[e_i|D_i = 1] - E[e_i|D_i = 0])$

Explaining the OVB

Q: Can you explain the OVB in our case? Can you guess a sign?

A: OVB: $(E[e_i|D_i = 1] - E[e_i|D_i = 0])$

- If the errors are related to AUM, then OVB exists

Explaining the OVB

Q: Can you explain the OVB in our case? Can you guess a sign?

A: OVB: $(E[e_i|D_i = 1] - E[e_i|D_i = 0])$

- If the errors are related to AUM, then OVB exists
- If the portion of revenues not explained by our model are different for the treat and control samples, then we have an OVB

Explaining the OVB

Q: Can you explain the OVB in our case? Can you guess a sign?

A: OVB: $(E[e_i|D_i = 1] - E[e_i|D_i = 0])$

- If the errors are related to AUM, then OVB exists
- If the portion of revenues not explained by our model are different for the treat and control samples, then we have an OVB
- Suppose the client suspects there might be something about the large firms that makes the more successful (i.e. higher Revenue) than the smaller firms

Explaining the OVB

Q: Can you explain the OVB in our case? Can you guess a sign?

A: OVB: $(E[e_i|D_i = 1] - E[e_i|D_i = 0])$

- If the errors are related to AUM, then OVB exists
- If the portion of revenues not explained by our model are different for the treat and control samples, then we have an OVB
- Suppose the client suspects there might be something about the large firms that makes the more successful (i.e. higher Revenue) than the smaller firms
- Q: Sign of OVB?

Explaining the OVB

Q: Can you explain the OVB in our case? Can you guess a sign?

A: OVB: $(E[e_i|D_i = 1] - E[e_i|D_i = 0])$

- If the errors are related to AUM, then OVB exists
- If the portion of revenues not explained by our model are different for the treat and control samples, then we have an OVB
- Suppose the client suspects there might be something about the large firms that makes the more successful (i.e. higher Revenue) than the smaller firms
- Q: Sign of OVB?
- A: positive

Focus on AUM

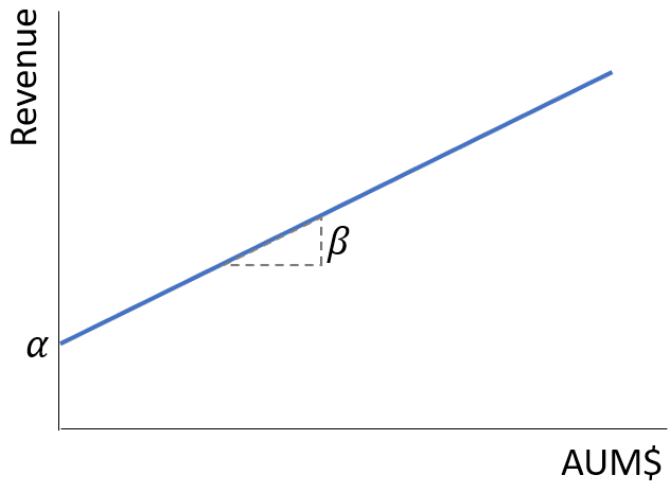
Q: Why not run a regression of Revenue on AUM?

Focus on AUM

Q: Why not run a regression of Revenue on AUM?

A:

$$Rev_i = \alpha + \beta AUM_i + e_i$$

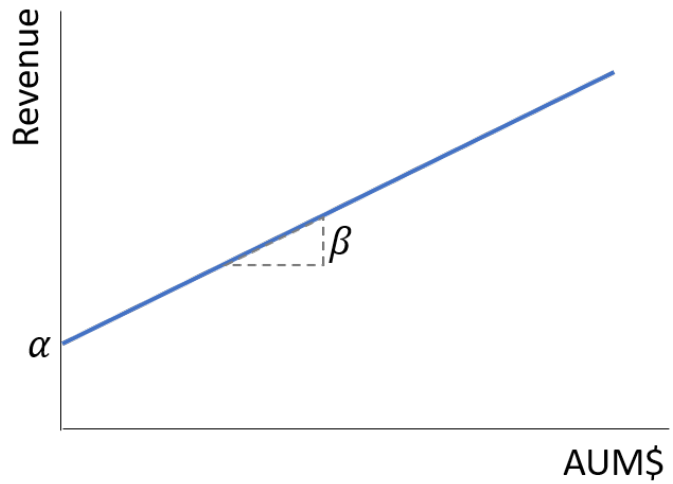


Focus on AUM

Q: Why not run a regression of Revenue on AUM?

A:

$$Rev_i = \alpha + \beta AUM_i + e_i$$



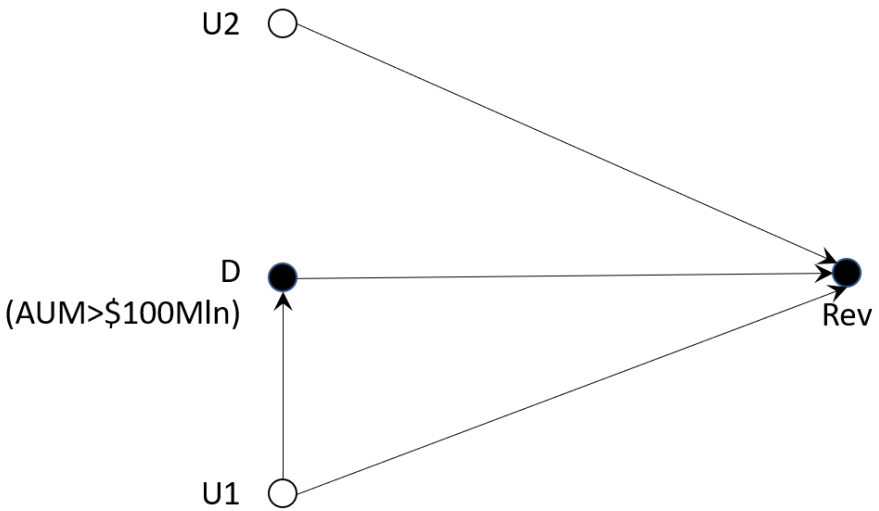
This doesn't reveal the TE of SEC regulation upon revenues. Rather, it shows the (constant) marginal effect of an increase in AUM on Revenue. We want to focus on the marginal effect in the neighborhood of $100mln.

Outline

1. Case Study
2. Incorrectly Estimating TE
3. Sharp RD
4. Fuzzy RD

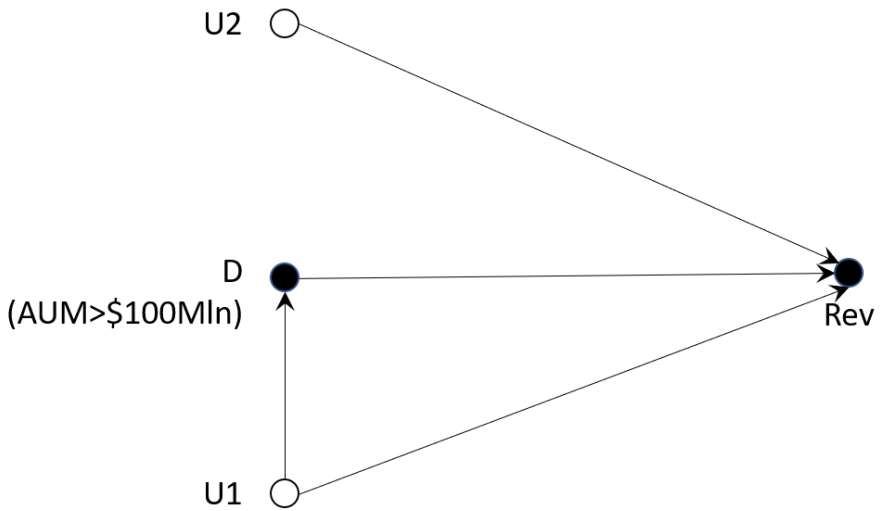
Extending the DAG

Q: What might be in U1 and U2?



Extending the DAG

Q: What might be in U1 and U2?

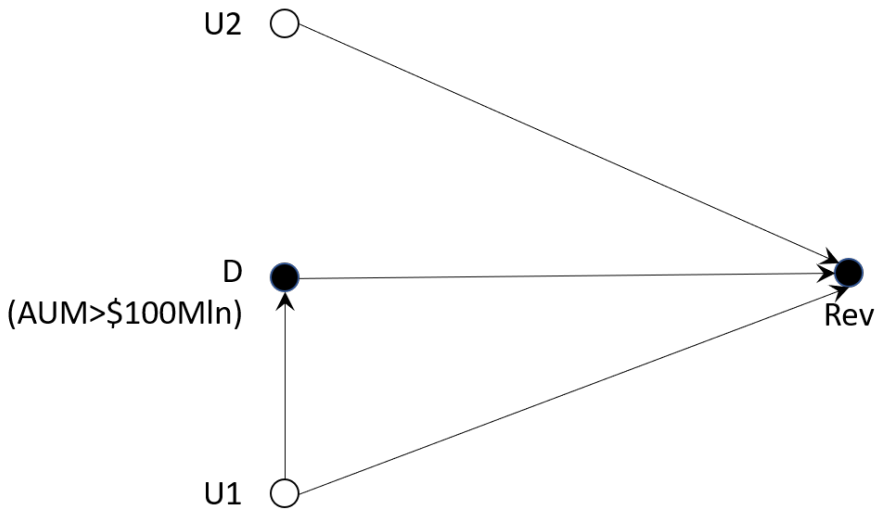


A:

- U1: Size of sales team, experience (e.g. number of transactions cleared), etc.. i.e. AUM related

Extending the DAG

Q: What might be in U1 and U2?

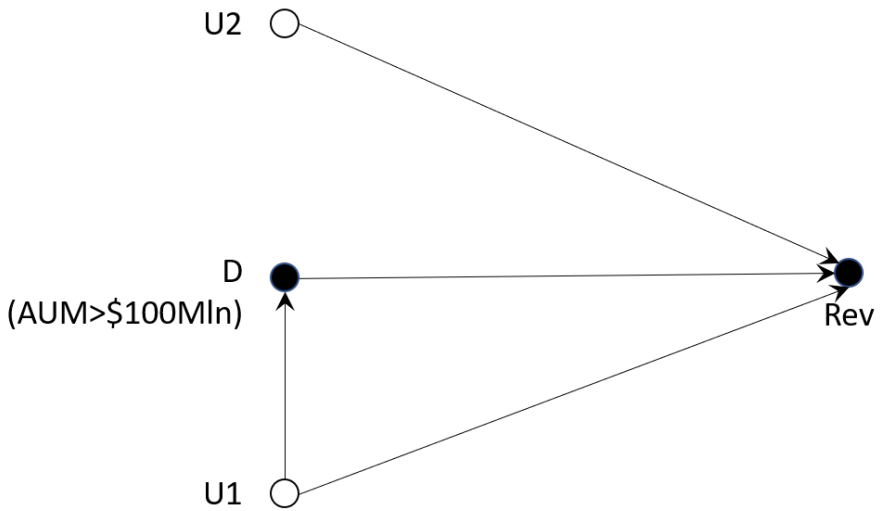


A:

- U1: Size of sales team, experience (e.g. number of transactions cleared), etc.. i.e. AUM related
- U2: Ancillary services that impact revenues but don't count towards AUM, such as flat fee financial planning

Extending the DAG

Q: What might be in U1 and U2?



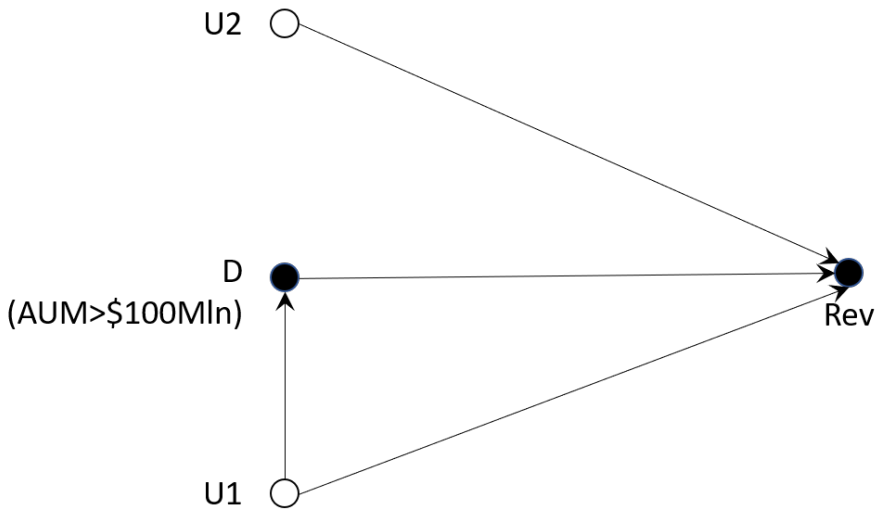
A:

- U1: Size of sales team, experience (e.g. number of transactions cleared), etc.. i.e. AUM related
- U2: Ancillary services that impact revenues but don't count towards AUM, such as flat fee financial planning

Q: Which is most important for estimation of the treatment effect?

Extending the DAG

Q: What might be in U1 and U2?



A:

- U1: Size of sales team, experience (e.g. number of transactions cleared), etc.. i.e. AUM related
- U2: Ancillary services that impact revenues but don't count towards AUM, such as flat fee financial planning

Q: Which is most important for estimation of the treatment effect?

A: U1 is a "Good Control". Including U2 might help with precision, but that isn't our focus at the moment.

Overcoming what's missing

Q: What can we do if we don't have data on the items in U1?

Overcoming what's missing

Q: What can we do if we don't have data on the items in U1?

A: We can include AUM directly.

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + u_i$$

Overcoming what's missing

Q: What can we do if we don't have data on the items in U1?

A: We can include AUM directly.

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + u_i$$

By including AUM in our model we are accommodating for all of the AUM related drivers in U1. In this case, AUM is referred to as the "running variable". Everything else is related to the error term.

Overcoming what's missing

Q: What can we do if we don't have data on the items in U1?

A: We can include AUM directly.

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + u_i$$

By including AUM in our model we are accommodating for all of the AUM related drivers in U1. In this case, AUM is referred to as the "running variable". Everything else is related to the error term.

Q: What is the RD estimate of the ATE?

Overcoming what's missing

Q: What can we do if we don't have data on the items in U1?

A: We can include AUM directly.

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + u_i$$

By including AUM in our model we are accommodating for all of the AUM related drivers in U1. In this case, AUM is referred to as the "running variable". Everything else is related to the error term.

Q: What is the RD estimate of the ATE?
A: β

Caution using full sample

Q: What's the danger in estimating this model on the full sample?

Caution using full sample

Q: What's the danger in estimating this model on the full sample?

A:

- As always, we want the treatment to be as close to randomly assigned as possible.

Caution using full sample

Q: What's the danger in estimating this model on the full sample?

A:

- As always, we want the treatment to be as close to randomly assigned as possible.
- Which means we want the State regulated firms to look like the SEC regulated firms.

Caution using full sample

Q: What's the danger in estimating this model on the full sample?

A:

- As always, we want the treatment to be as close to randomly assigned as possible.
- Which means we want the State regulated firms to look like the SEC regulated firms.
- But very large firms might behave differently than very small firms.

Caution using full sample

Q: What's the danger in estimating this model on the full sample?

A:

- As always, we want the treatment to be as close to randomly assigned as possible.
- Which means we want the State regulated firms to look like the SEC regulated firms.
- But very large firms might behave differently than very small firms.
- The Sharp RD solution is

Caution using full sample

Q: What's the danger in estimating this model on the full sample?

A:

- As always, we want the treatment to be as close to randomly assigned as possible.
- Which means we want the State regulated firms to look like the SEC regulated firms.
- But very large firms might behave differently than very small firms.
- The Sharp RD solution is to impose a caliper

The Sharp RD Caliper

Q: What's the logic behind a sharp RD caliper?

The Sharp RD Caliper

Q: What's the logic behind a sharp RD caliper?

A:

- Comparing a $\$100K$ RIA to a $\$100Bln$ is probably not fair.

The Sharp RD Caliper

Q: What's the logic behind a sharp RD caliper?

A:

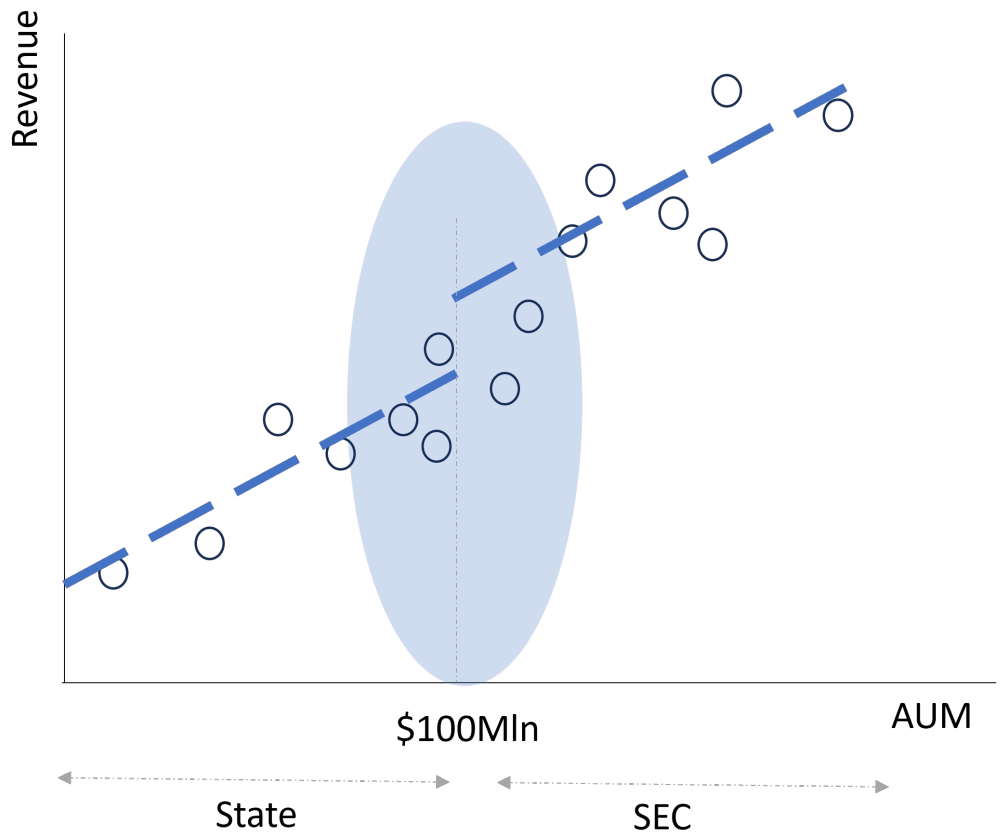
- Comparing a $\$100K$ RIA to a $\$100Bln$ is probably not fair.
- However, a $\$99Mln$ RIA is probably very similar to a $\$101Mln$ in every way except for regulatory coverage.

The Sharp RD Caliper

Q: What's the logic behind a sharp RD caliper?

A:

- Comparing a $\$100K$ RIA to a $\$100Bln$ is probably not fair.
- However, a $\$99Mln$ RIA is probably very similar to a $\$101Mln$ in every way except for regulatory coverage.



Setting the Caliper

Q: What's the tradeoff with setting the caliper size?

Setting the Caliper

Q: What's the tradeoff with setting the caliper size?
A:

- Larger Caliper: more data, but

Setting the Caliper

Q: What's the tradeoff with setting the caliper size?
A:

- Larger Caliper: more data, but influenced by extraneous forces

Setting the Caliper

Q: What's the tradeoff with setting the caliper size?
A:

- Larger Caliper: more data, but influenced by extraneous forces
- Smaller Caliper: sharper estimates, but

Setting the Caliper

Q: What's the tradeoff with setting the caliper size?
A:

- Larger Caliper: more data, but influenced by extraneous forces
- Smaller Caliper: sharper estimates, but less data implies less confidence in results

Setting the Caliper

Q: What's the tradeoff with setting the caliper size?
A:

- Larger Caliper: more data, but influenced by extraneous forces
- Smaller Caliper: sharper estimates, but less data implies less confidence in results

Best Practices

Setting the Caliper

Q: What's the tradeoff with setting the caliper size?
A:

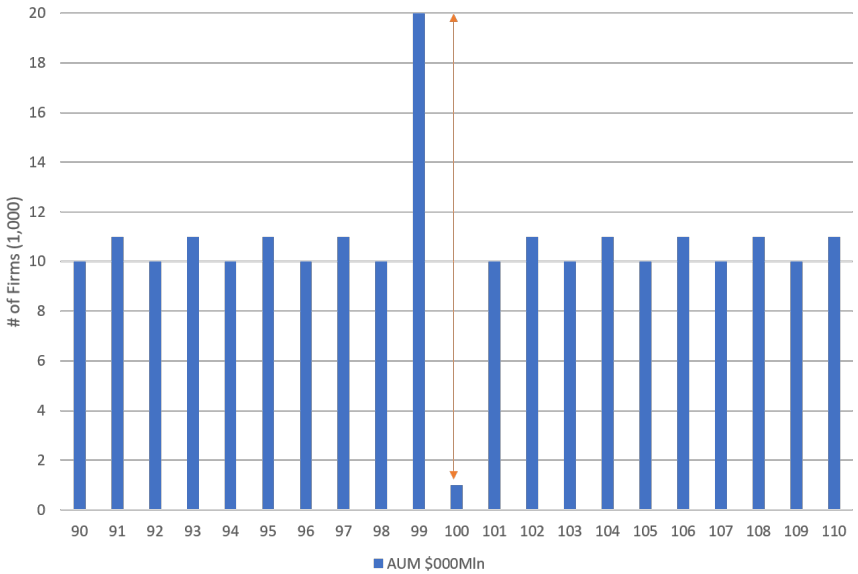
- Larger Caliper: more data, but influenced by extraneous forces
- Smaller Caliper: sharper estimates, but less data implies less confidence in results

Best Practices

- Try a few caliper sizes
- Start narrow with reasonable sample size
- Increase caliper slowly
- Stop when estimates stabilize

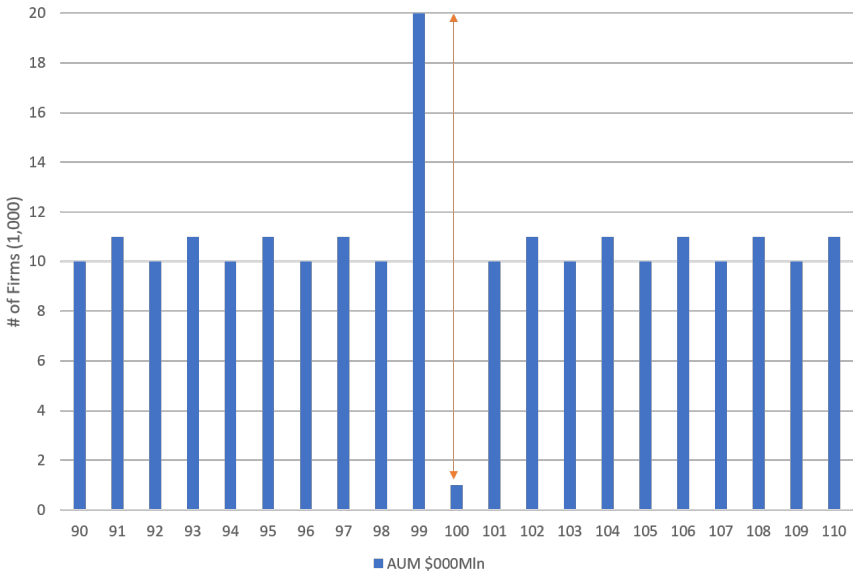
Watch out for manipulation of treatment

Q: Why are these sample statistics a concern for RD?



Watch out for manipulation of treatment

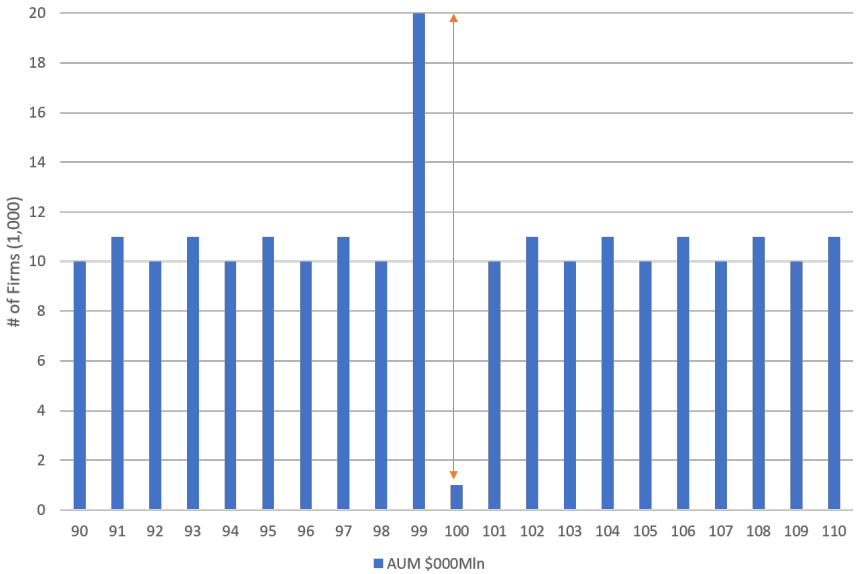
Q: Why are these sample statistics a concern for RD?



A: Suppose there are two types of firms. Type A are all revenue via AUM. Type B are a mix of AUM and other services. Assume only firms with a mix of services (Type B) could restrict their AUM while maintaining profitability.

Watch out for manipulation of treatment

Q: Why are these sample statistics a concern for RD?

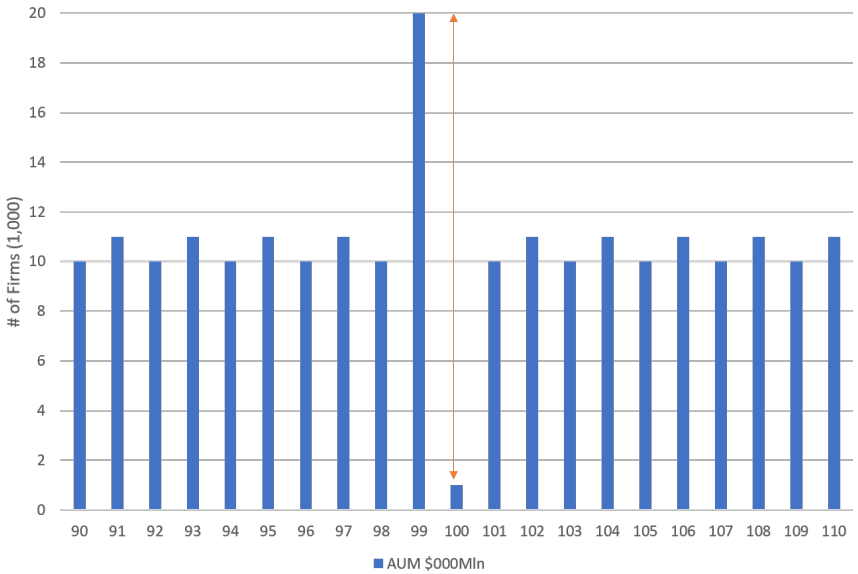


A: Suppose there are two types of firms. Type A are all revenue via AUM. Type B are a mix of AUM and other services. Assume only firms with a mix of services (Type B) could restrict their AUM while maintaining profitability.

- At $\$99 Mln$ we have Type A and Type B firms. While at $\$100 Mln$ we have only Type A firms.

Watch out for manipulation of treatment

Q: Why are these sample statistics a concern for RD?



A: Suppose there are two types of firms. Type A are all revenue via AUM. Type B are a mix of AUM and other services. Assume only firms with a mix of services (Type B) could restrict their AUM while maintaining profitability.

- At $\$99 \mathit{Mln}$ we have Type A and Type B firms. While at $\$100 \mathit{Mln}$ we have only Type A firms.
- The RD design is invalid. The Type B have manipulated their treatment by intentionally not growing their AUM so as to avoid the SEC regulatory burden.

Sharp RD Summary

Requirements for Sharp RD

- Treatment is a deterministic function of running variable
- Treatment is discontinuous function of running variable
- No manipulation of treatment

Sharp RD Summary

Requirements for Sharp RD

- Treatment is a deterministic function of running variable
- Treatment is discontinuous function of running variable
- No manipulation of treatment

Steps for Conducting Sharp RD

1. Identify the outcome, treatment, and running variables

Sharp RD Summary

Requirements for Sharp RD

- Treatment is a deterministic function of running variable
- Treatment is discontinuous function of running variable
- No manipulation of treatment

Steps for Conducting Sharp RD

1. Identify the outcome, treatment, and running variables
2. Create a treatment dummy from the running variable

Sharp RD Summary

Requirements for Sharp RD

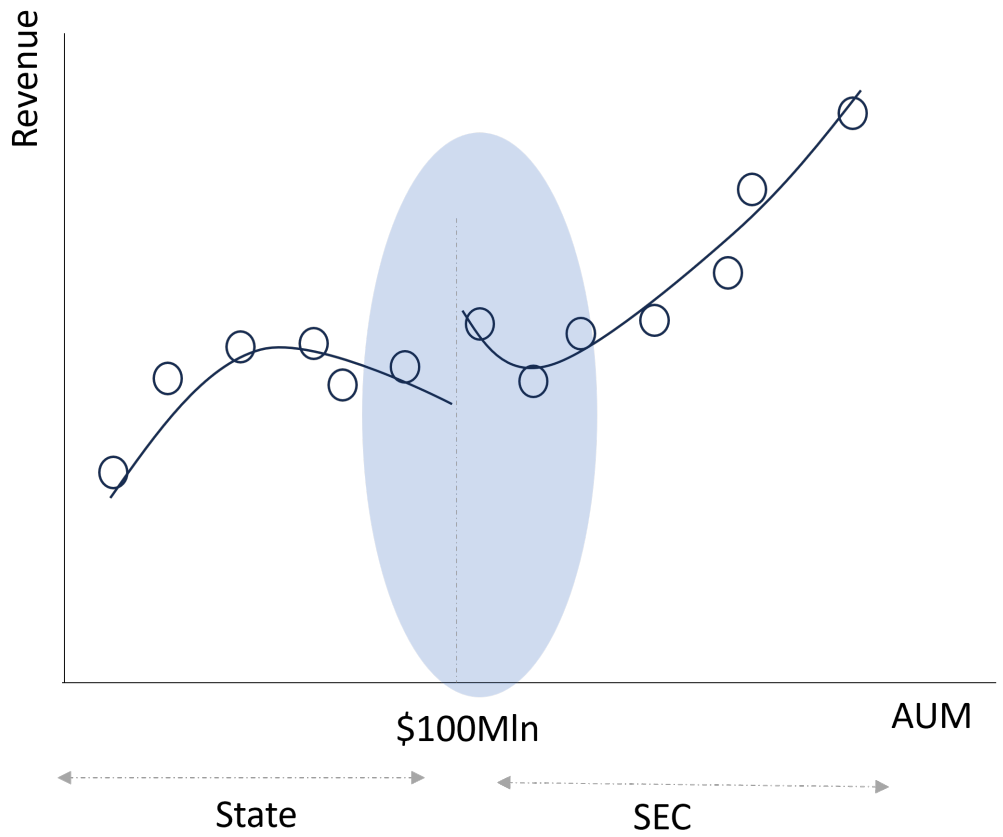
- Treatment is a deterministic function of running variable
- Treatment is discontinuous function of running variable
- No manipulation of treatment

Steps for Conducting Sharp RD

1. Identify the outcome, treatment, and running variables
2. Create a treatment dummy from the running variable
3. Regress outcome on treatment and running variable

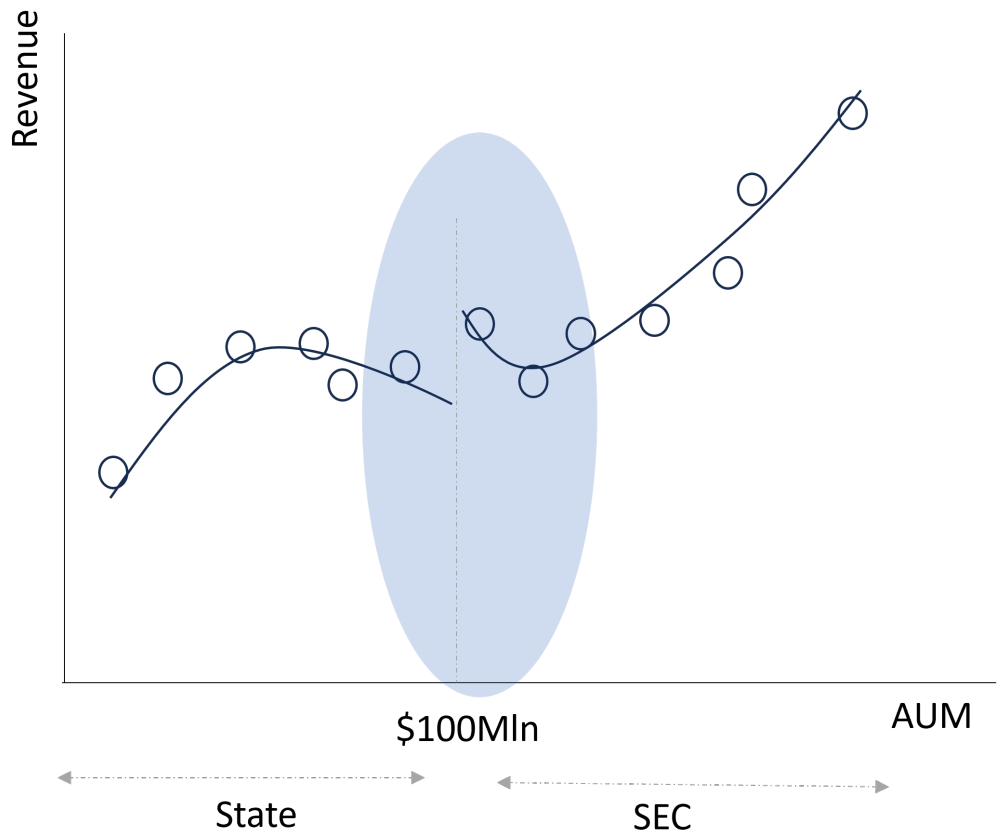
Nonlinear Running Variable Effects

Q: How accommodate nonlinear effects?



Nonlinear Running Variable Effects

Q: How accommodate nonlinear effects?

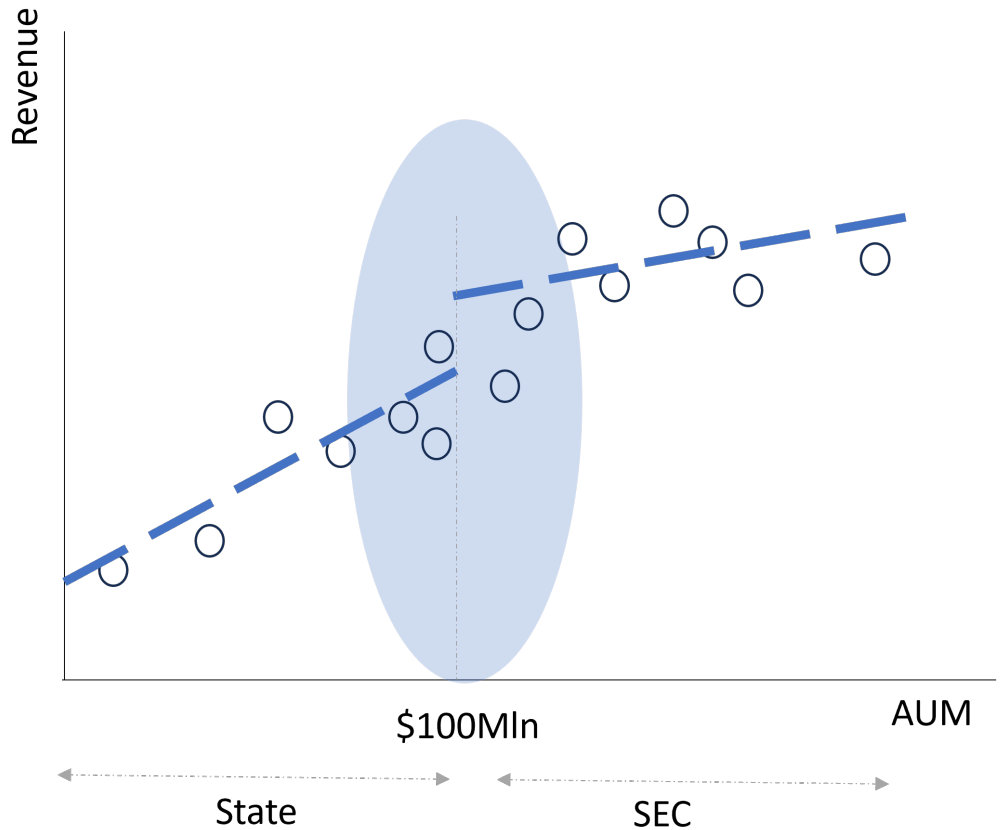


A:

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta AUM^2 + \cdots + u_i$$

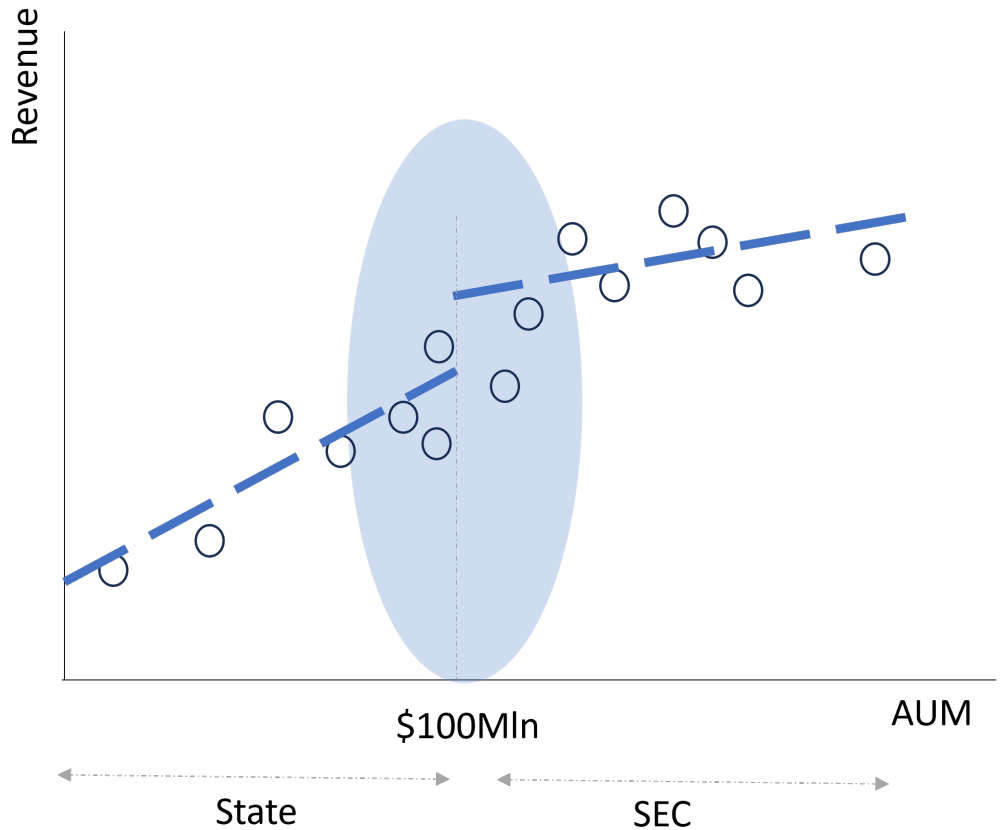
Differing Running Variable Effects

Q: How accommodate differing marginal effects of the running variable on either side of the threshold?



Differing Running Variable Effects

Q: How accommodate differing marginal effects of the running variable on either side of the threshold?



$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$E[Rev|D = 1, AUM] =$$

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma AUM_i + \delta AUM_i$$

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma AUM_i + \delta AUM_i$$
$$E[Rev|D = 0, AUM] =$$

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma AUM_i + \delta AUM_i$$
$$E[Rev|D = 0, AUM] = \alpha + \gamma AUM_i$$

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma AUM_i + \delta AUM_i$$
$$E[Rev|D = 0, AUM] = \alpha + \gamma AUM_i$$
$$ATE =$$

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma AUM_i + \delta AUM_i$$
$$E[Rev|D = 0, AUM] = \alpha + \gamma AUM_i$$
$$ATE = \beta + \delta AUM_i$$

Difficulty with Interpretation

Q: How represent the ATE at the \$100*Mln* threshold in the prior model?

$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$\begin{aligned} E[Rev|D = 1, AUM] &= \alpha + \beta + \gamma AUM_i + \delta AUM_i \\ E[Rev|D = 0, AUM] &= \alpha + \gamma AUM_i \\ ATE &= \beta + \delta AUM_i \\ ATE|100Mln &= \end{aligned}$$

Difficulty with Interpretation

Q: How represent the ATE at the $\$100 Mln$ threshold in the prior model?

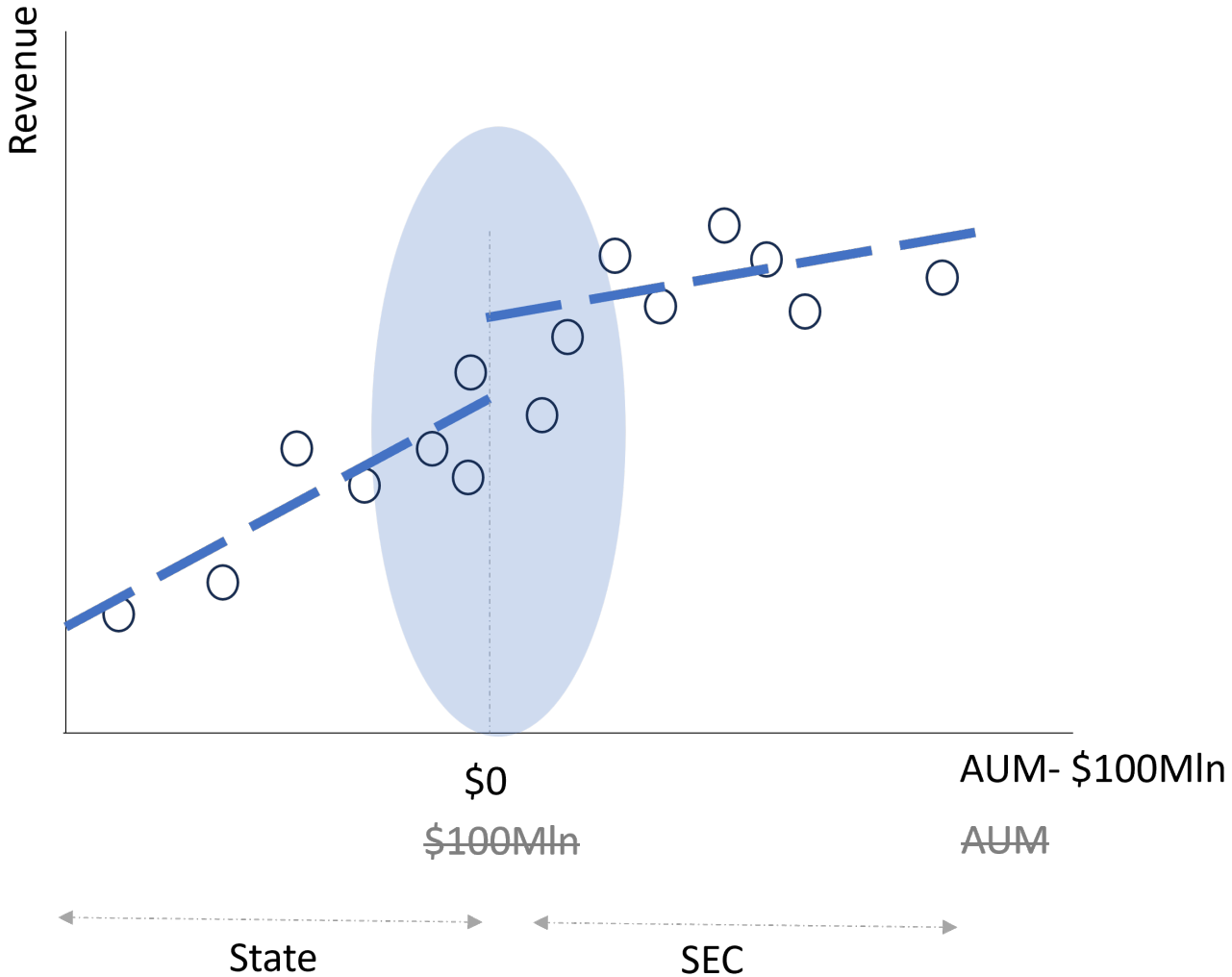
$$Rev_i = \alpha + \beta D_i + \gamma AUM_i + \delta D_i \times AUM + u_i$$

A:

$$\begin{aligned} E[Rev|D = 1, AUM] &= \alpha + \beta + \gamma AUM_i + \delta AUM_i \\ E[Rev|D = 0, AUM] &= \alpha + \gamma AUM_i \\ ATE &= \beta + \delta AUM_i \\ ATE|100Mln &= \beta + \delta \times \$100Mln \end{aligned}$$

Redefining the Running Variable

For ease of interpretation, we can center the running variable



Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at \$100*Mln*?

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$E[Rev|D = 1, AUM] =$

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma(AUM_i - 100Mln) + \delta(AUM_i - 100Mln)$$

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma(AUM_i - 100Mln) + \delta(AUM_i - 100Mln)$$
$$E[Rev|D = 0, AUM] =$$

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma(AUM_i - 100Mln) + \delta(AUM_i - 100Mln)$$
$$E[Rev|D = 0, AUM] = \alpha + \gamma(AUM_i - 100Mln)$$

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma(AUM_i - 100Mln) + \delta(AUM_i - 100Mln)$$
$$E[Rev|D = 0, AUM] = \alpha + \gamma(AUM_i - 100Mln)$$
$$ATE =$$

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$$E[Rev|D = 1, AUM] = \alpha + \beta + \gamma(AUM_i - 100Mln) + \delta(AUM_i - 100Mln)$$
$$E[Rev|D = 0, AUM] = \alpha + \gamma(AUM_i - 100Mln)$$
$$ATE = \beta + \delta(AUM_i - 100Mln)$$

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$$\begin{aligned} E[Rev|D = 1, AUM] &= \alpha + \beta + \gamma(AUM_i - 100Mln) + \delta(AUM_i - 100Mln) \\ E[Rev|D = 0, AUM] &= \alpha + \gamma(AUM_i - 100Mln) \\ ATE &= \beta + \delta(AUM_i - 100Mln) \\ ATE|100Mln &= \end{aligned}$$

Easier Interpretation

$$Rev_i = \alpha + \beta D_i + \gamma(AUM_i - 100Mln) + \delta D_i \times (AUM - 100Mln) + u_i$$

Q: How interpret the ATE at $\$100Mln$?
A:

$$\begin{aligned} E[Rev|D = 1, AUM] &= \alpha + \beta + \gamma(AUM_i - 100Mln) + \delta(AUM_i - 100Mln) \\ E[Rev|D = 0, AUM] &= \alpha + \gamma(AUM_i - 100Mln) \\ ATE &= \beta + \delta(AUM_i - 100Mln) \\ ATE|100Mln &= \beta \end{aligned}$$

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)
- Suppose there are two tiers of CRM service: business and institutional

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)
- Suppose there are two tiers of CRM service: business and institutional
- Suppose the institutional version is only sold to firms with AUM $\geq \$100 Mln$?

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)
- Suppose there are two tiers of CRM service: business and institutional
- Suppose the institutional version is only sold to firms with AUM $\geq \$100 Mln$?
- Q: What is the trouble we face in computing our ATE?

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)
- Suppose there are two tiers of CRM service: business and institutional
- Suppose the institutional version is only sold to firms with AUM $\geq \$100 Mln$?
- Q: What is the trouble we face in computing our ATE?
- A: We can't disentangle the "SEC Effect" from the "CRM Effect"

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)
- Suppose there are two tiers of CRM service: business and institutional
- Suppose the institutional version is only sold to firms with AUM $\geq$ \$100*Mln*?
- Q: What is the trouble we face in computing our ATE?
- A: We can't disentangle the "SEC Effect" from the "CRM Effect"
- We can only uncover an impact of the \$100*Mln* threshold

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)
- Suppose there are two tiers of CRM service: business and institutional
- Suppose the institutional version is only sold to firms with AUM $\geq \$100 Mln$?
- Q: What is the trouble we face in computing our ATE?
- A: We can't disentangle the "SEC Effect" from the "CRM Effect"
- We can only uncover an impact of the $\$100 Mln$ threshold
- Likewise, Sharp RD cannot disentangle the SEC Effect from a discrete "Big Firm Effect."

Overlapping Treatments

- RIA's are big users of CRM tools (Customer Relationship Management software)
- Suppose there are two tiers of CRM service: business and institutional
- Suppose the institutional version is only sold to firms with AUM $\geq \$100 Mln$?
- Q: What is the trouble we face in computing our ATE?
- A: We can't disentangle the "SEC Effect" from the "CRM Effect"
- We can only uncover an impact of the $\$100 Mln$ threshold
- Likewise, Sharp RD cannot disentangle the SEC Effect from a discrete "Big Firm Effect."
- But, under the right conditions, Fuzzy RD may be able to isolate the SEC effect.

Outline

1. Case Study
2. Incorrectly Estimating TE
3. Sharp RD
4. Fuzzy RD

A Fuzzy Treatment

Suppose the regulator is not assigned via a deterministic rule

- SEC regulation is based upon AUM $\geq \$100 Mln +$ "Other"

A Fuzzy Treatment

Suppose the regulator is not assigned via a deterministic rule

- SEC regulation is based upon AUM $\geq \$100 Mln$ + "Other"
- "Other" may include non-AUM items such as the amount of insurance sales or the relative strictness of a given State's regulatory burden

A Fuzzy Treatment

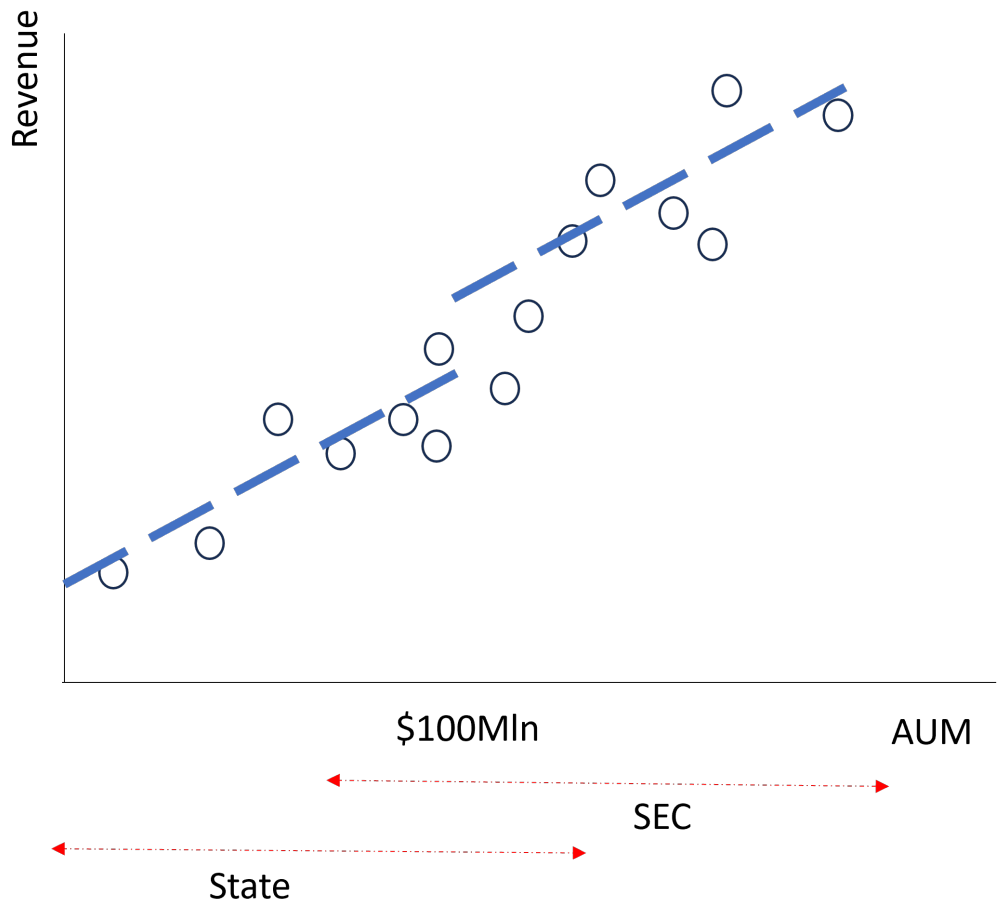
Suppose the regulator is not assigned via a deterministic rule

- SEC regulation is based upon AUM $\geq \$100 Mln$ + "Other"
- "Other" may include non-AUM items such as the amount of insurance sales or the relative strictness of a given State's regulatory burden
- We now have Fuzzy RD

A Fuzzy Treatment

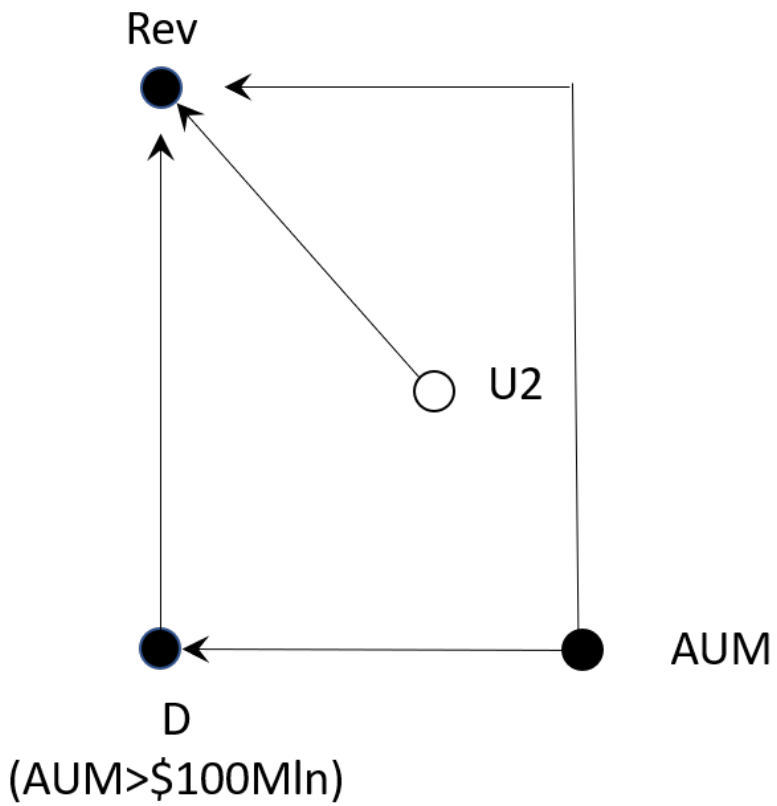
Suppose the regulator is not assigned via a deterministic rule

- SEC regulation is based upon AUM $\geq$ \$100*Mln* + "Other"
- "Other" may include non-AUM items such as the amount of insurance sales or the relative strictness of a given State's regulatory burden
- We now have Fuzzy RD



Spotting Trouble in the DAG

Our treatment variable D_{AUM} is no longer a sharp indication of treatment status



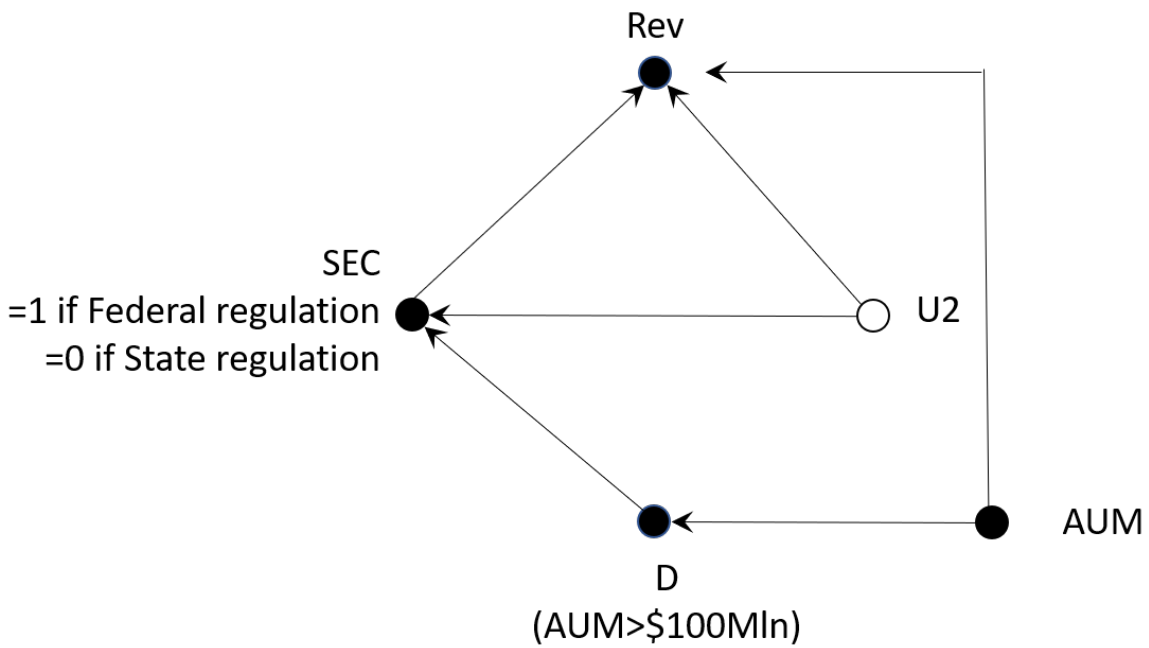
Slightly rearranged DAG to accommodate an important extension

Finding a solution in the DAG

Define SEC=1 to indicate SEC oversight

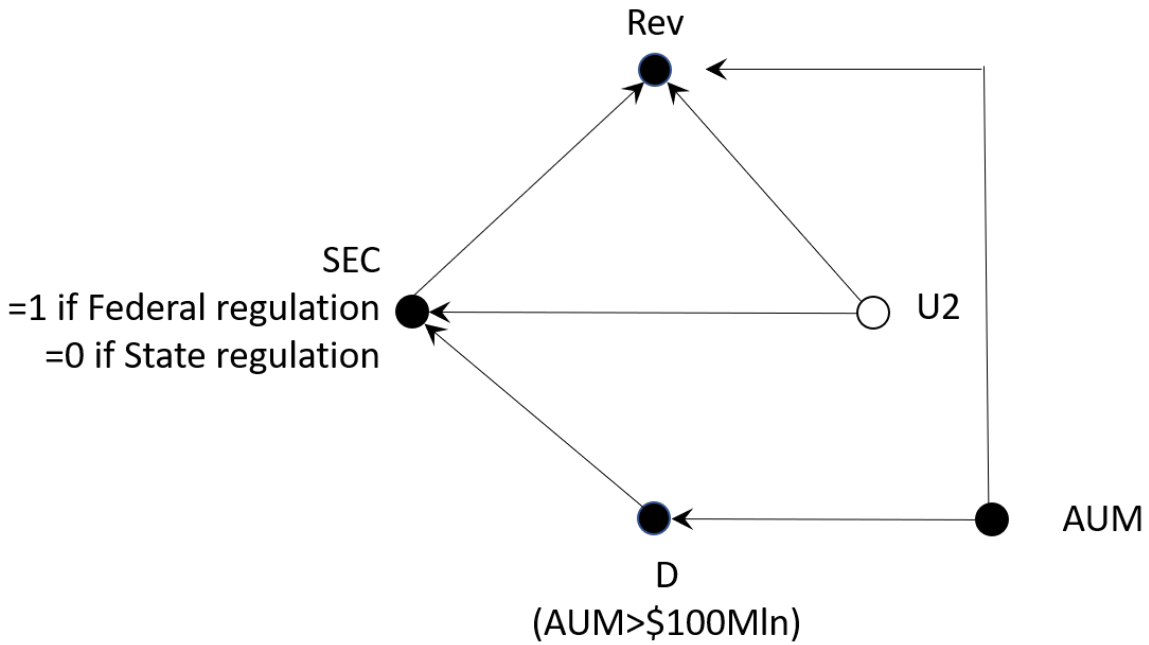
Finding a solution in the DAG

Define SEC=1 to indicate SEC oversight



Finding a solution in the DAG

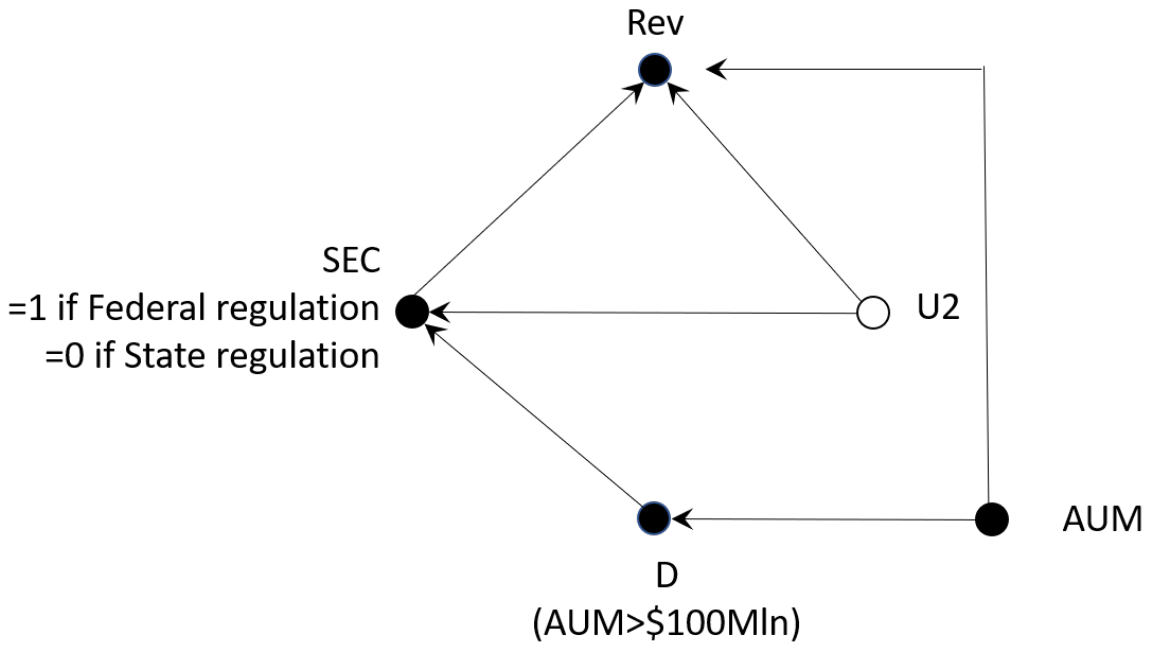
Define SEC=1 to indicate SEC oversight



- Q: Which arrow is the ATE?

Finding a solution in the DAG

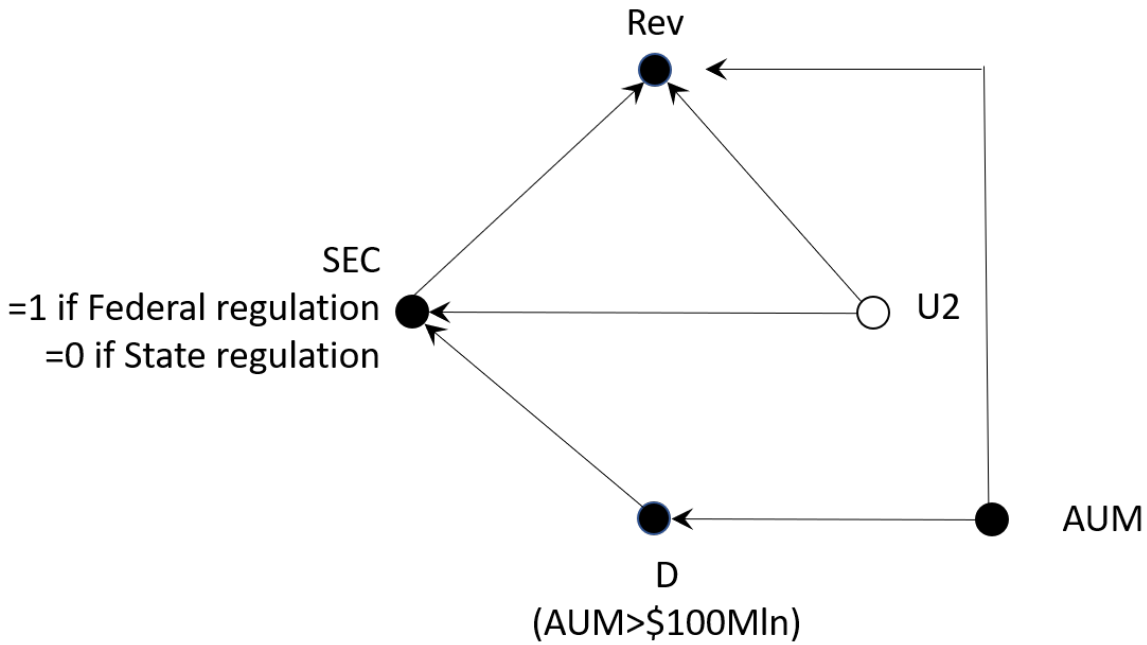
Define SEC=1 to indicate SEC oversight



- Q: Which arrow is the ATE? A: connection between SEC and Rev

Finding a solution in the DAG

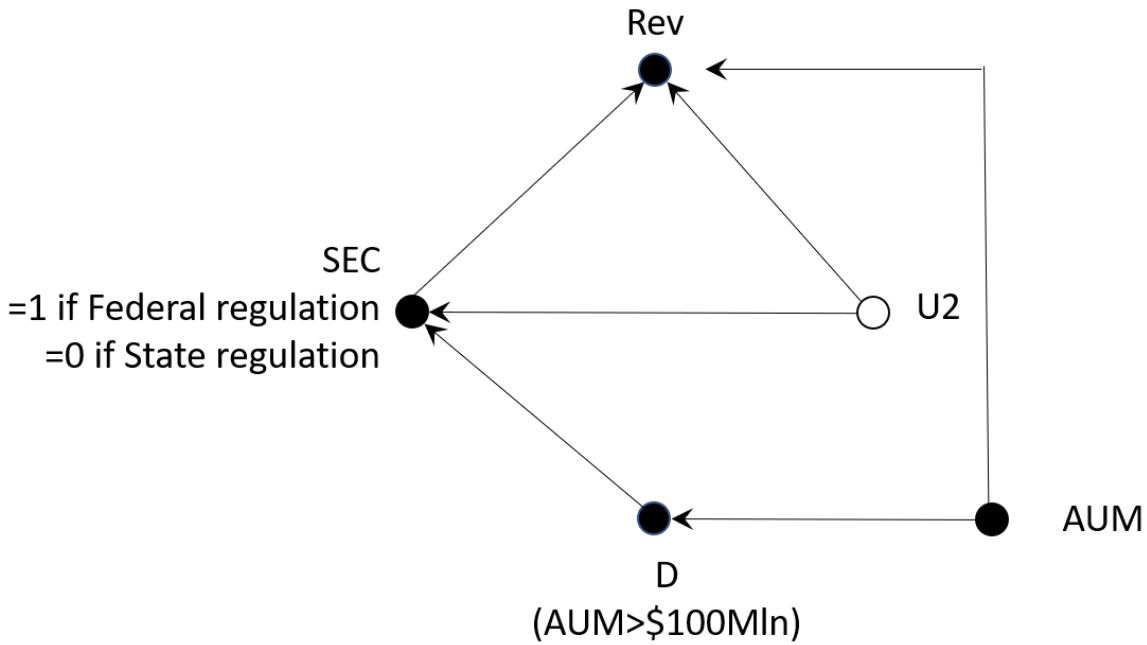
Define SEC=1 to indicate SEC oversight



- Q: Which arrow is the ATE? A: connection between SEC and Rev
- Q: Which variable creates an OVB?

Finding a solution in the DAG

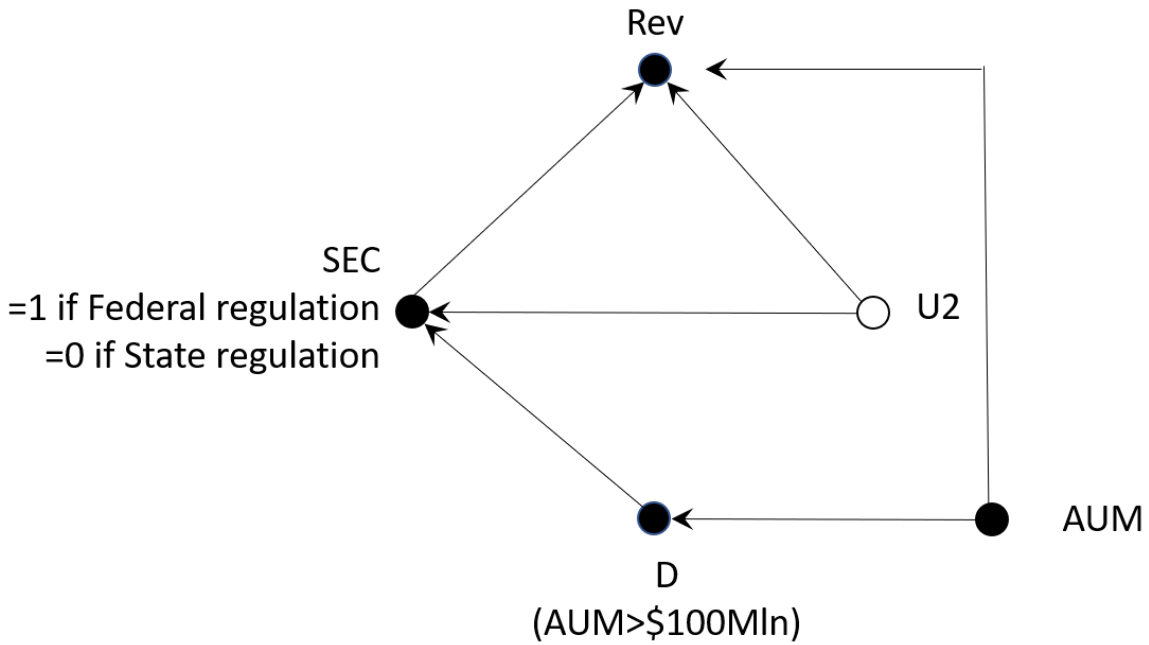
Define SEC=1 to indicate SEC oversight



- Q: Which arrow is the ATE? A: connection between SEC and Rev
- Q: Which variable creates an OVB? A: U2

Finding a solution in the DAG

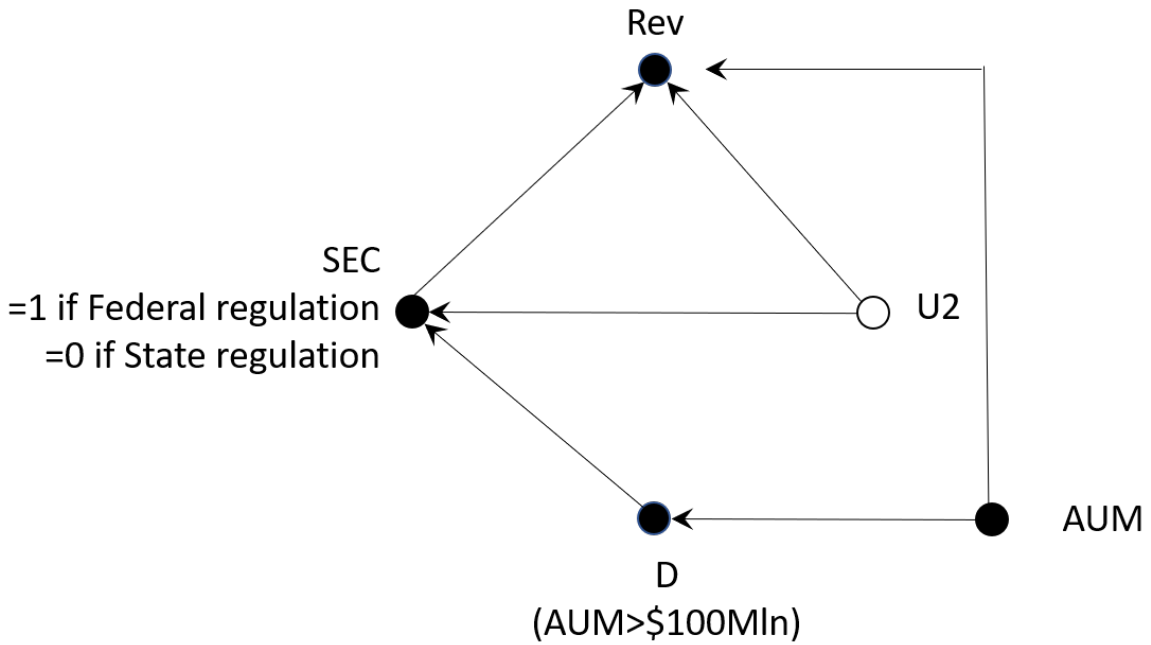
Define SEC=1 to indicate SEC oversight



- Q: Which arrow is the ATE? A: connection between SEC and Rev
- Q: Which variable creates an OVB? A: U2
- Q: Which of these variables can be used as an instrument for SEC?

Finding a solution in the DAG

Define SEC=1 to indicate SEC oversight



- Q: Which arrow is the ATE? A: connection between SEC and Rev
- Q: Which variable creates an OVB? A: U2
- Q: Which of these variables can be used as an instrument for SEC?
- A: D_{AUM}

IV Solution

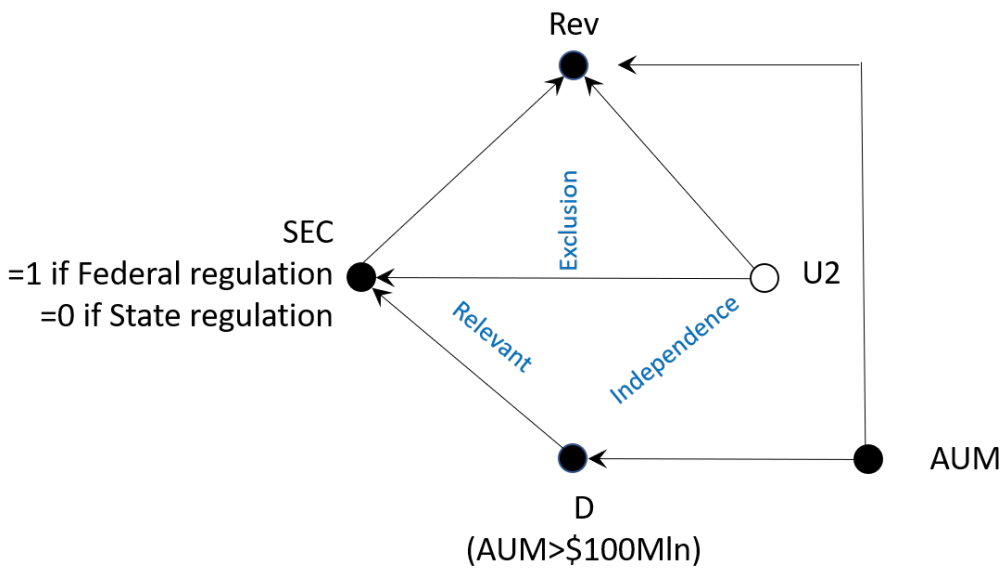
Q: Is D_{AUM} a valid instrument? Why/Why Not?

IV Solution

Q: Is D_{AUM} a valid instrument? Why/Why Not?

A: Yes

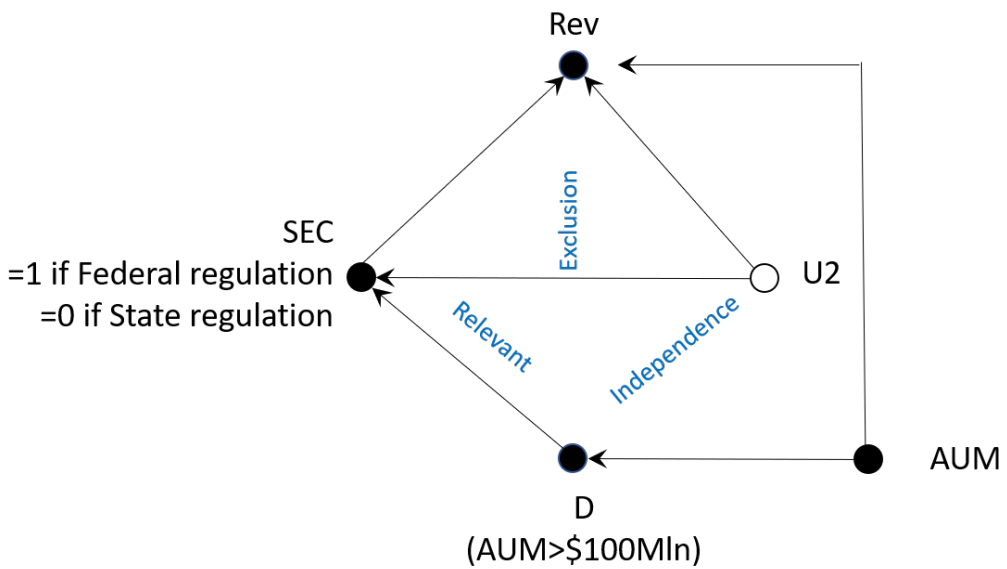
- Relevant:



IV Solution

Q: Is D_{AUM} a valid instrument? Why/Why Not?

A: Yes

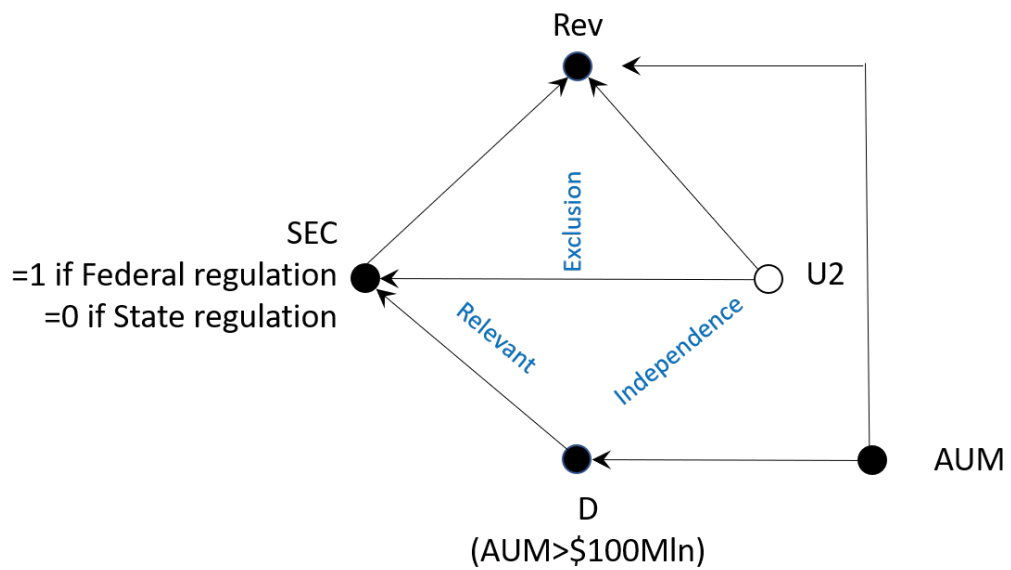


- Relevant: D is related to SEC via the (fuzzy) regulatory assignment rule

IV Solution

Q: Is D_{AUM} a valid instrument? Why/Why Not?

A: Yes

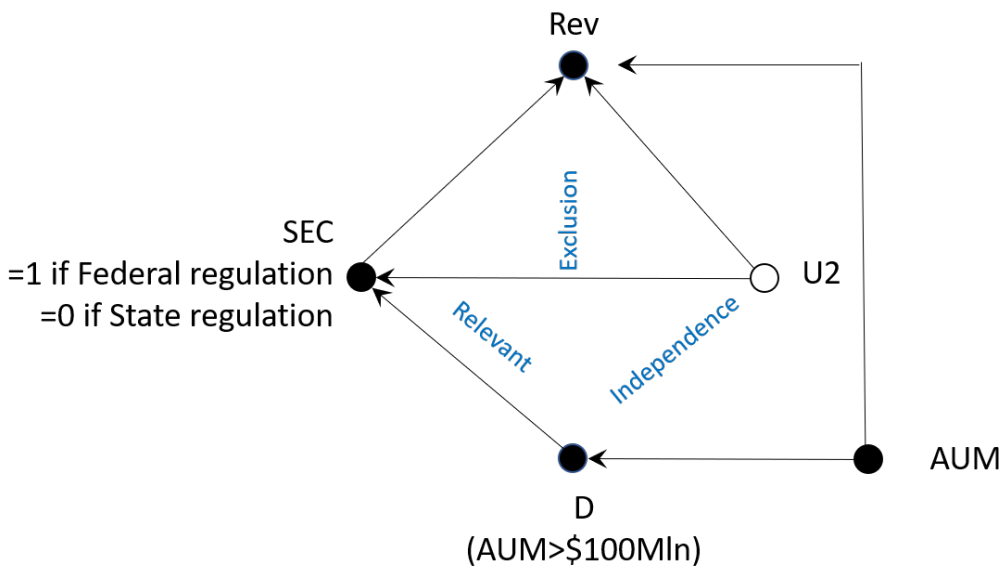


- Relevant: D is related to SEC via the (fuzzy) regulatory assignment rule
- Independence:

IV Solution

Q: Is D_{AUM} a valid instrument? Why/Why Not?

A: Yes

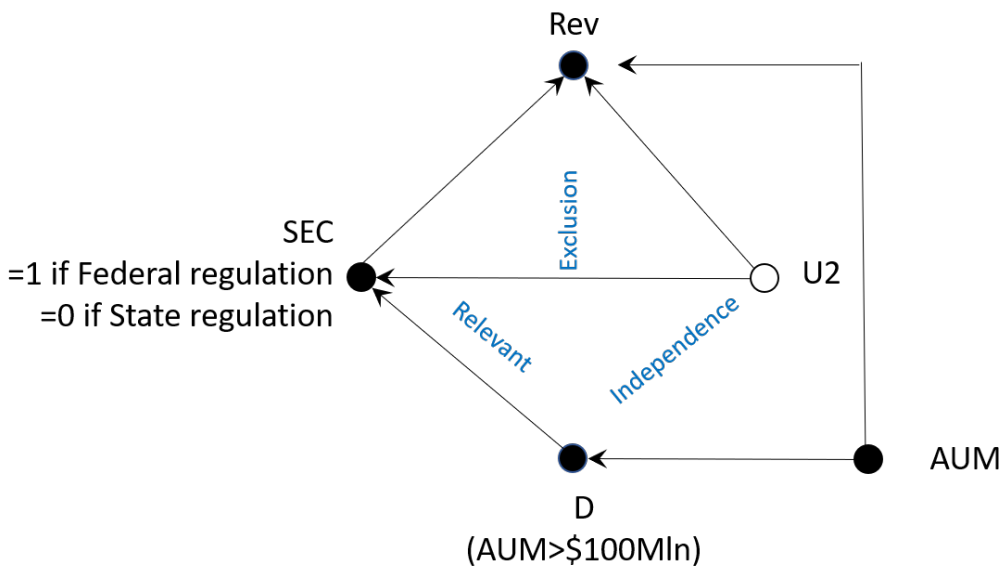


- Relevant: D is related to SEC via the (fuzzy) regulatory assignment rule
- Independence: D_{AUM} is not related to the errors (i.e. other drivers of revenues)
- Exclusion:

IV Solution

Q: Is D_{AUM} a valid instrument? Why/Why Not?

A: Yes



- Relevant: D is related to SEC via the (fuzzy) regulatory assignment rule
- Independence: D_{AUM} is not related to the errors (i.e. other drivers of revenues)
- Exclusion: D_{AUM} is not related to Revenues directly; only through SEC

IV avoids OVB with 2 steps

The first stage regression estimates SEC using AUM

$$SEC = \pi_0 + \pi_1 D_{AUM} + \pi_2 AUM + e$$

IV avoids OVB with 2 steps

The first stage regression estimates SEC using AUM

$$SEC = \pi_0 + \pi_1 D_{AUM} + \pi_2 AUM + e$$

The second stage regression uses that part of SEC that is related only to AUM (AUM and D_{AUM})

$$Rev = \alpha + \beta \widehat{SEC} + \gamma AUM + e$$

IV avoids OVB with 2 steps

The first stage regression estimates SEC using AUM

$$SEC = \pi_0 + \pi_1 D_{AUM} + \pi_2 AUM + e$$

The second stage regression uses that part of SEC that is related only to AUM (AUM and D_{AUM})

$$Rev = \alpha + \beta \widehat{SEC} + \gamma AUM + e$$

β now provides the marginal effect of a change in regulator that only has to do with AUM, thereby avoiding the non-AUM related drivers of SEC regulation that prompted the OVB.

With a discrete treatment, we could estimate the first stage via a binary model (e.g. Logit).

LATE

- Reminder: IV produces the LATE

LATE

- Reminder: IV produces the LATE
- Q: To whom does the IV apply in our case?

LATE

- Reminder: IV produces the LATE
- Q: To whom does the IV apply in our case?
- A:
  - IV produces the ATE for compliers

LATE

- Reminder: IV produces the LATE
- Q: To whom does the IV apply in our case?
- A:
  - IV produces the ATE for compliers
  - The compliers are the RIAs that follow the rules, i.e.

LATE

- Reminder: IV produces the LATE
- Q: To whom does the IV apply in our case?
- A:
  - IV produces the ATE for compliers
  - The compliers are the RIAs that follow the rules, i.e. if they are assigned SEC regulation, they take it, and if assigned State regulation, they take it.

LATE

- Reminder: IV produces the LATE
- Q: To whom does the IV apply in our case?
- A:
  - IV produces the ATE for compliers
  - The compliers are the RIAs that follow the rules, i.e. if they are assigned SEC regulation, they take it, and if assigned State regulation, they take it.
  - Our estimates don't apply to RIA's that

LATE

- Reminder: IV produces the LATE
- Q: To whom does the IV apply in our case?
- A:
  - IV produces the ATE for compliers
  - The compliers are the RIAs that follow the rules, i.e. if they are assigned SEC regulation, they take it, and if assigned State regulation, they take it.
  - Our estimates don't apply to RIA's that never follow the rules by always taking SEC regulation, never taking SEC regulation, or always doing the opposite of what they are assigned.

Regression Discontinuity - Summary

- Natural Experiment
- If the treatment is assigned according to a rule, then Regression Discontinuity (RD) might be able to reveal the ATE
- Sharp RD
  - Treatment is a deterministic, discontinuous function of "running" variable
  - Regress outcome on treatment and running variable
- Fuzzy RD
  - Treatment is assigned by more than just the running variable
  - Using a threshold dummy as an instrument for the treatment dummy generates an unbiased IV estimate of the ATE