Causal Inference with Treatment Measurement Error

A Non-parametric Instrumental Variable Approach

Yuchen Zhu, Limor Gultchin, Arthur Gretton, Matt Kusner, Ricardo Silva

Proceedings of the 38th Conference on Uncertainty in Artificial Intelligence | Oral presentation

Motivation

Motivation

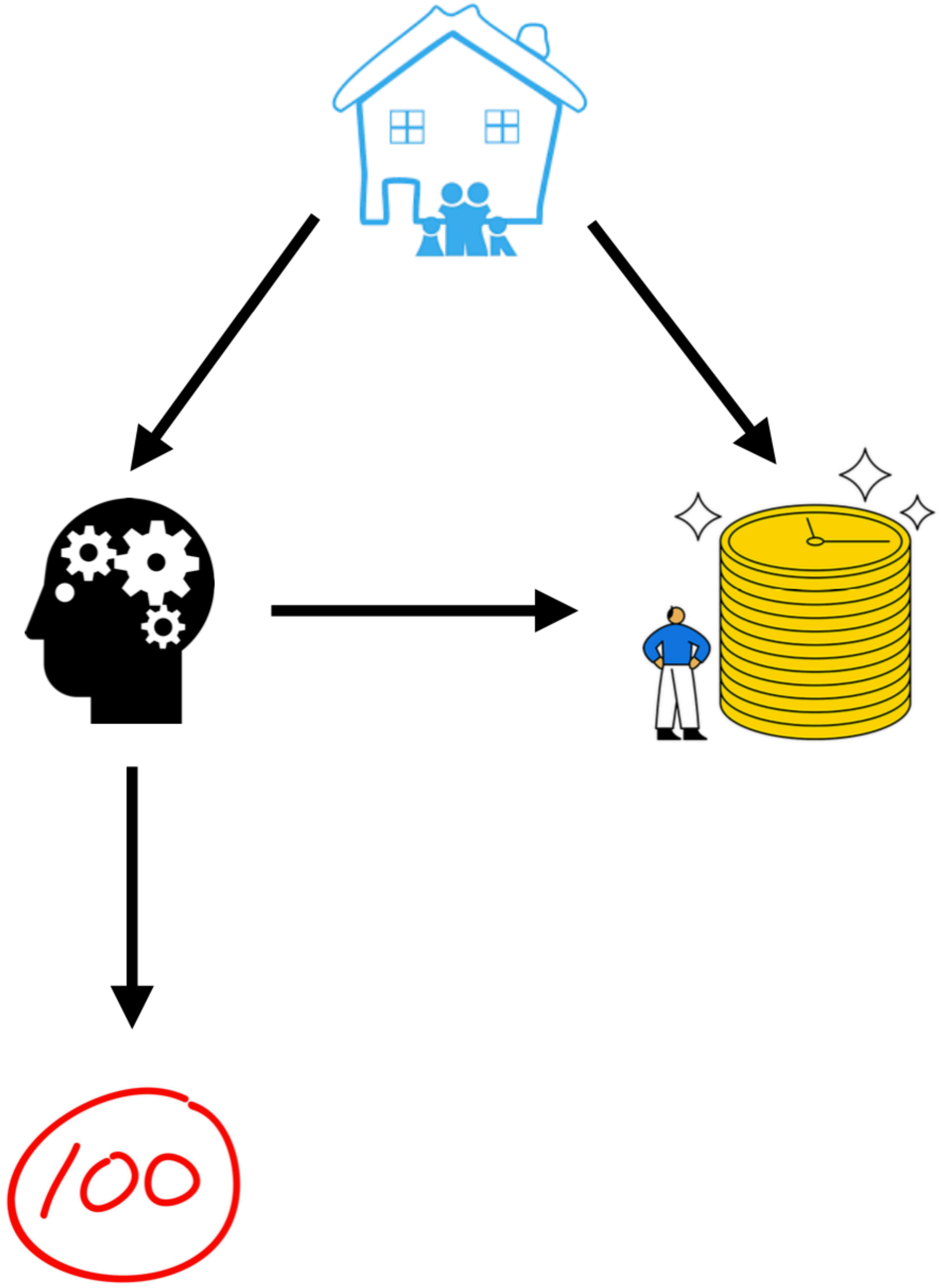
- Interesting causal questions often relates to notions imprecisely measured.

Motivation

- **Interesting causal questions often relates to notions imprecisely measured.**
  - **E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.**

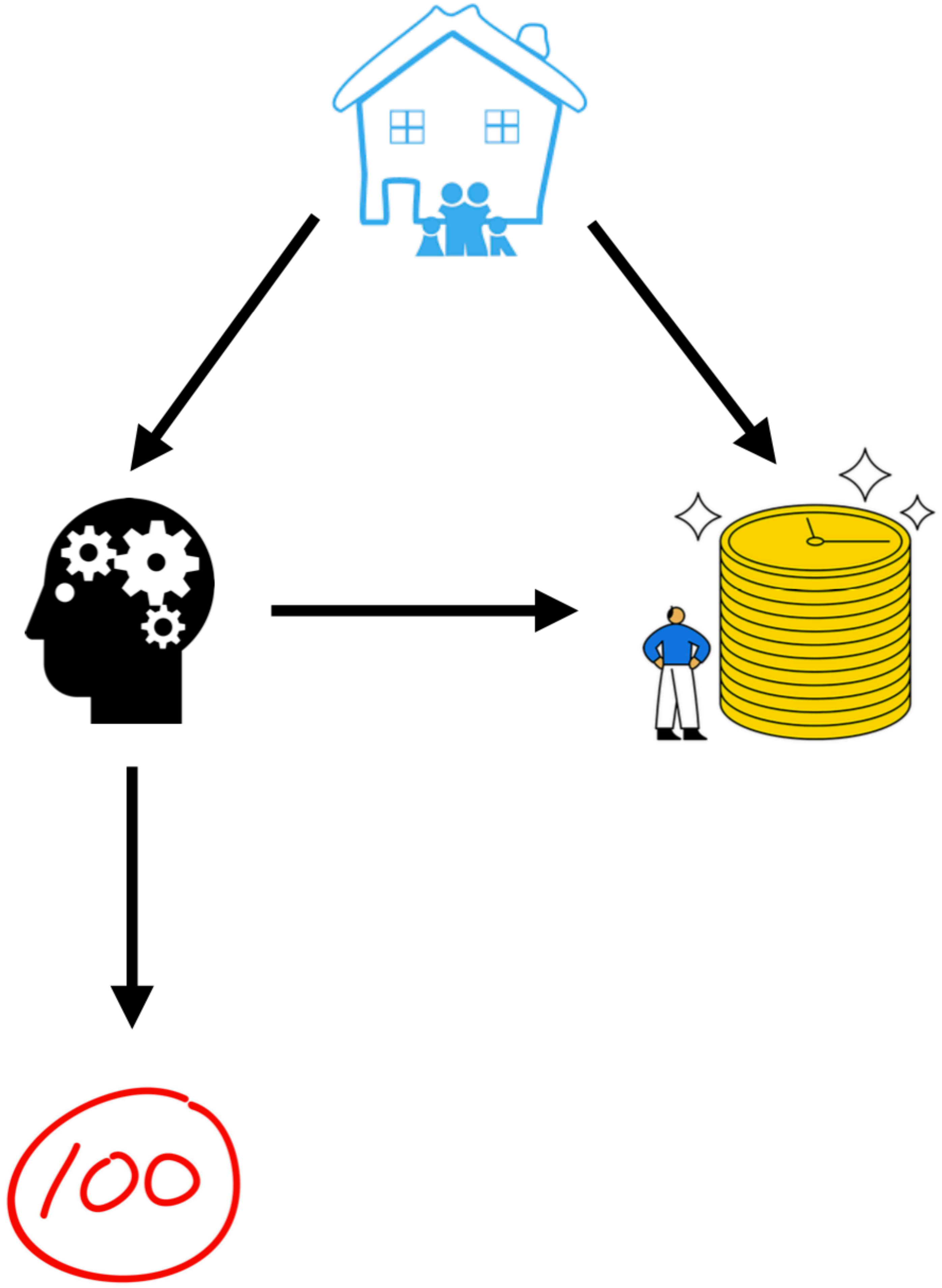
Motivation

- Interesting causal questions often relates to notions imprecisely measured.
  - E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.



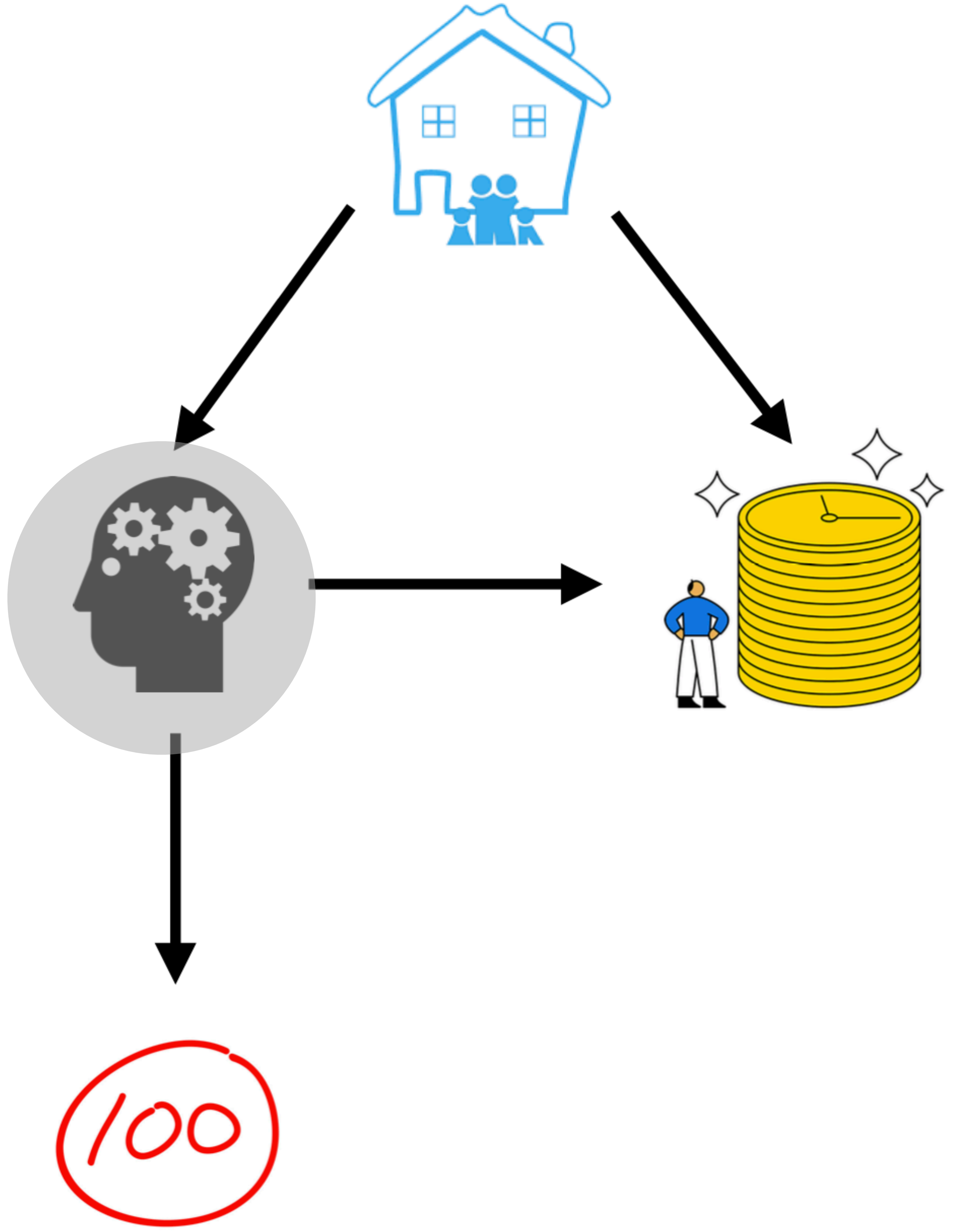
Motivation

- Interesting causal questions often relates to notions imprecisely measured.
  - E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.
- Two of the obstacles:



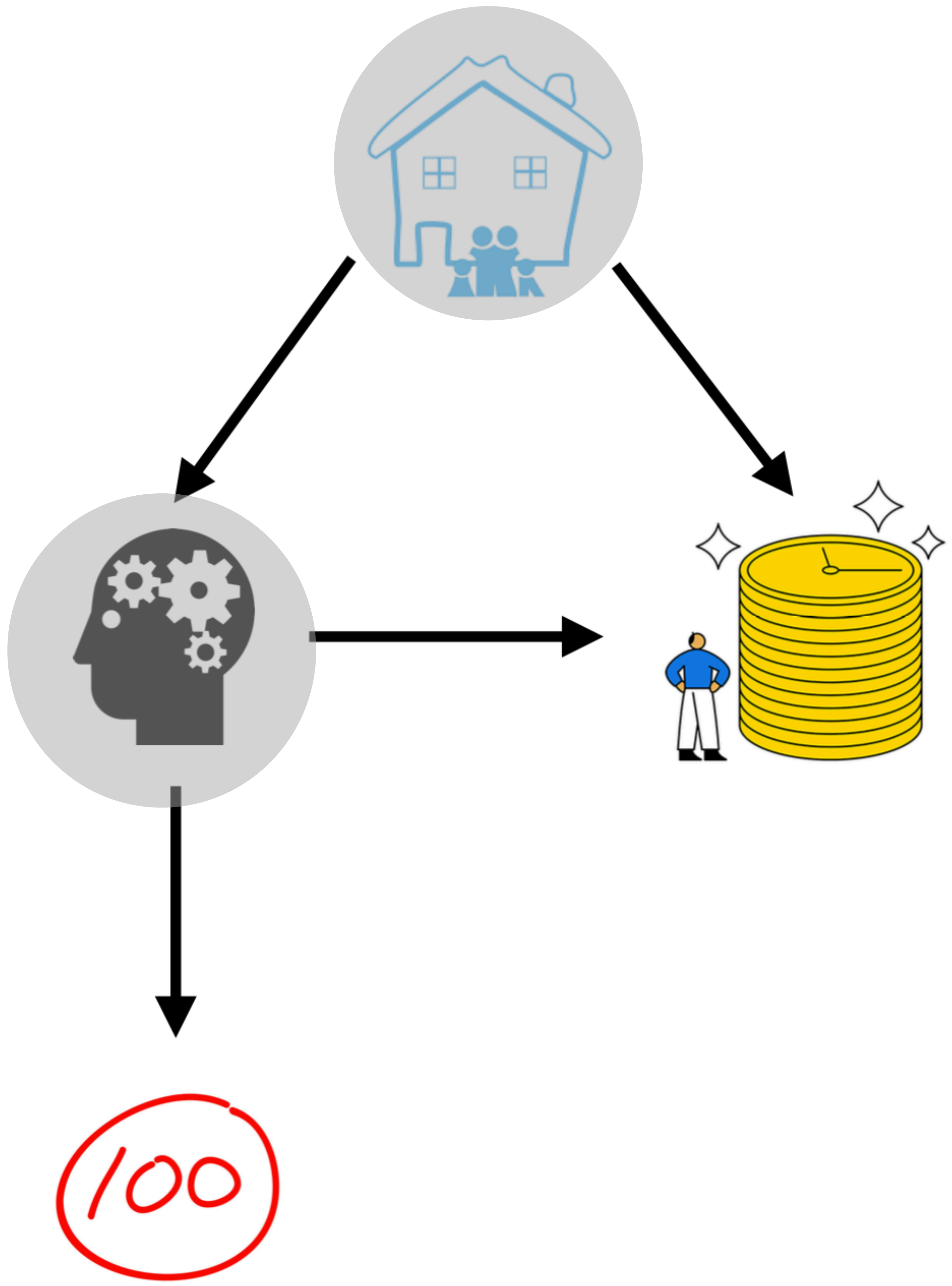
Motivation

- Interesting causal questions often relates to notions imprecisely measured.
  - E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.
- Two of the obstacles:
  - Measurement error on treatment.



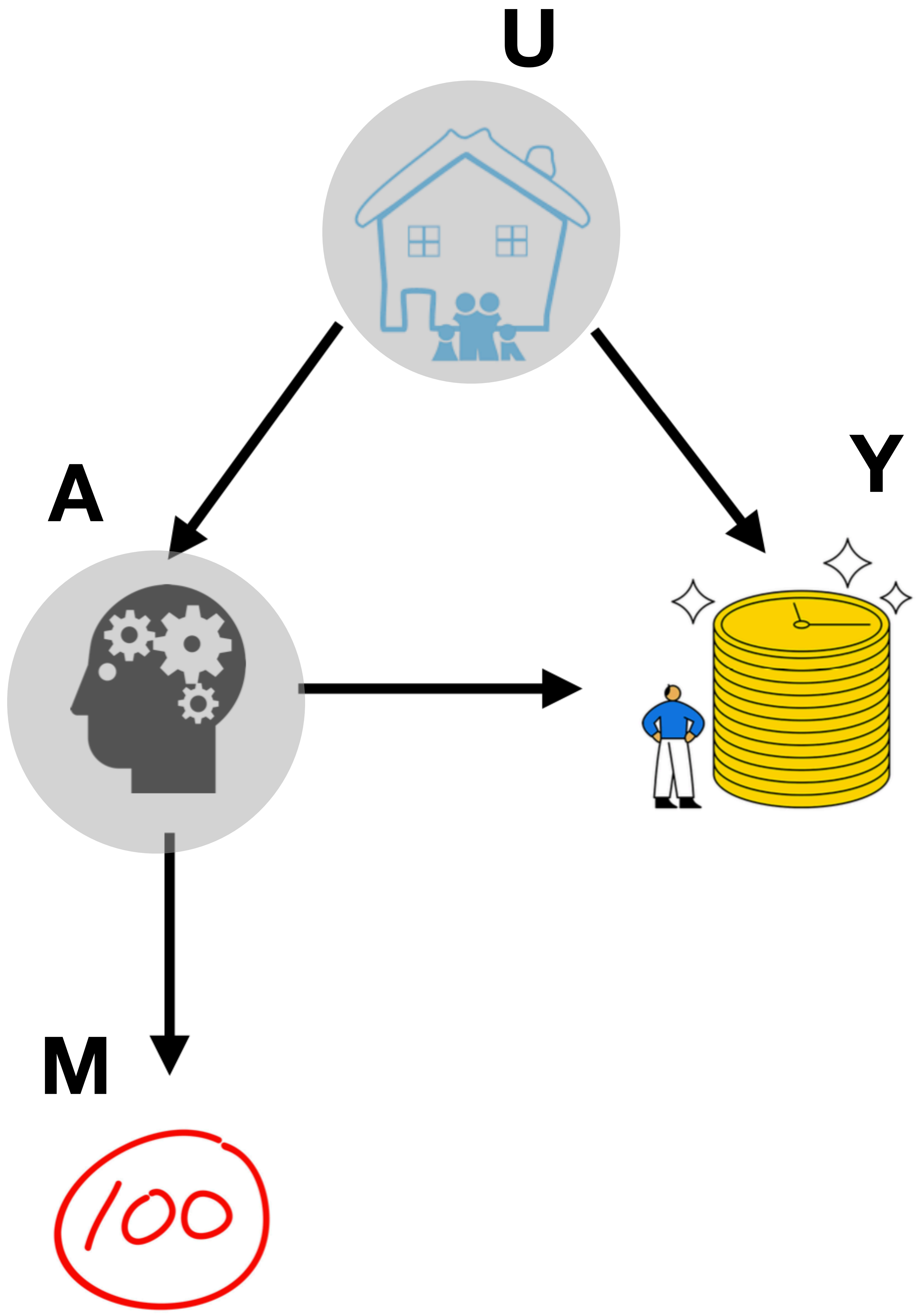
Motivation

- Interesting causal questions often relates to notions imprecisely measured.
  - E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.
- Two of the obstacles:
  - Measurement error on treatment.
  - Unobserved confounders.



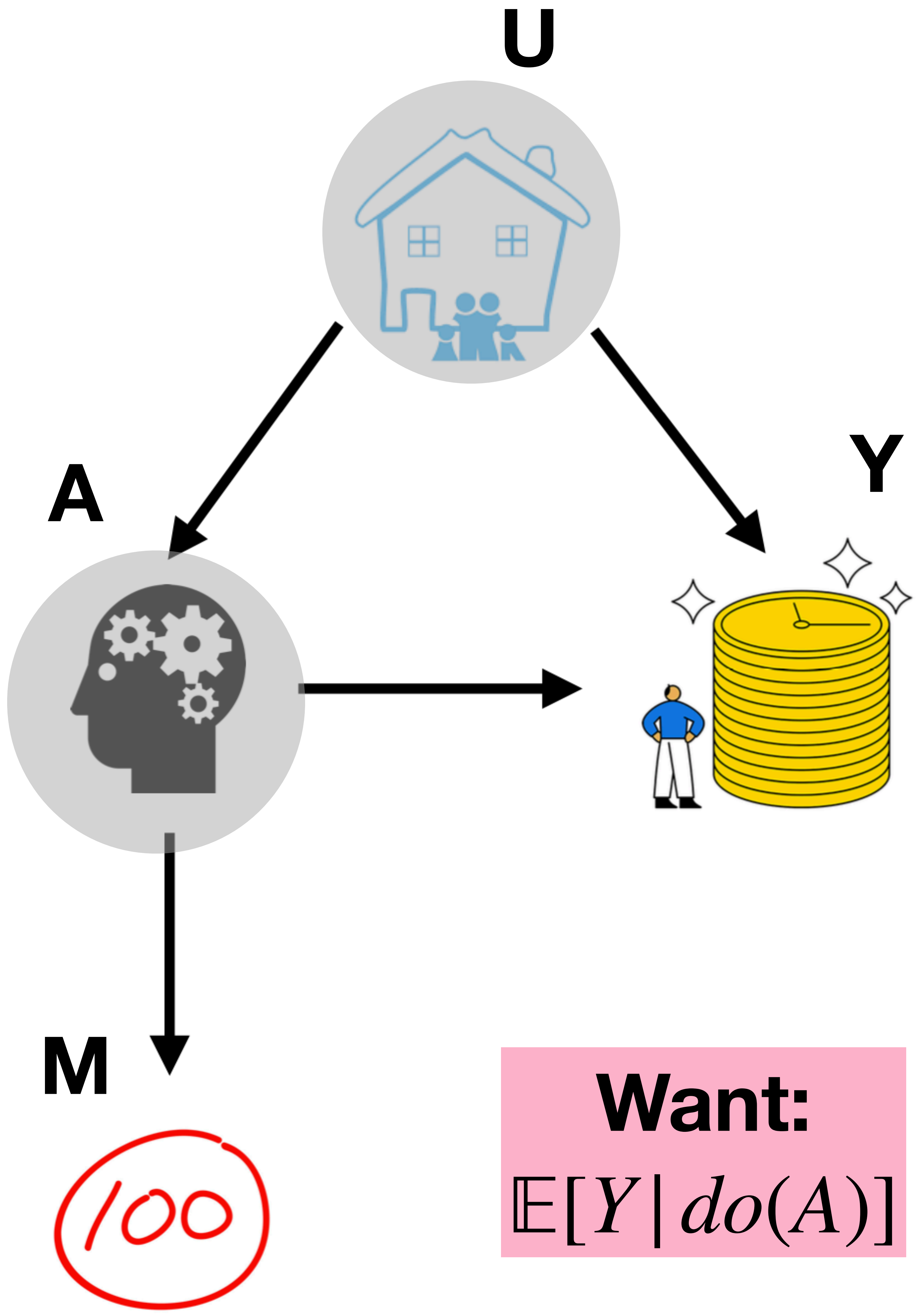
Motivation

- Interesting causal questions often relates to notions imprecisely measured.
  - E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.
- Two of the obstacles:
  - Measurement error on treatment.
  - Unobserved confounders.



Motivation

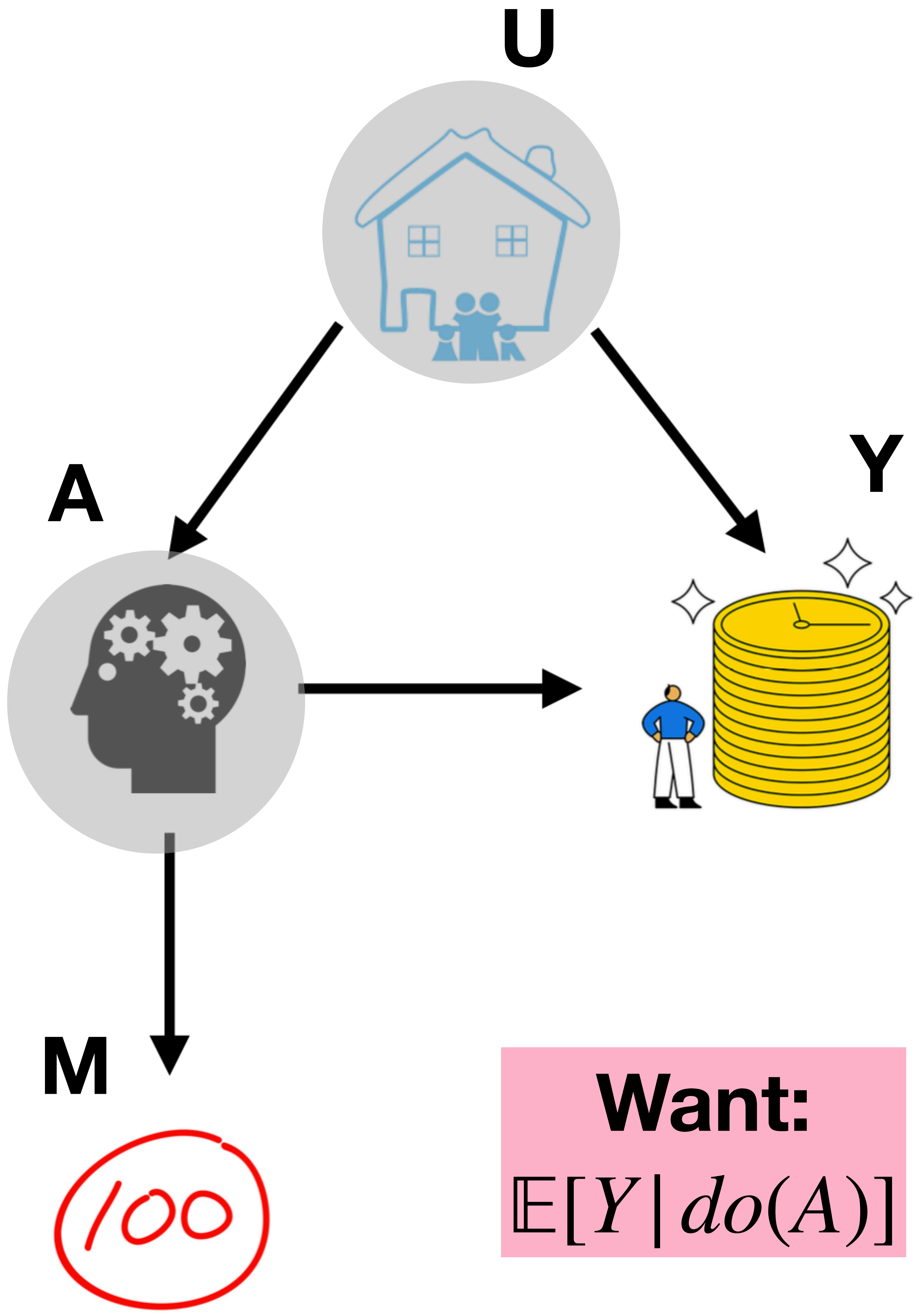
- Interesting causal questions often relates to notions imprecisely measured.
  - E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.
- Two of the obstacles:
  - Measurement error on treatment.
  - Unobserved confounders.



Want:
$\mathbb{E}[Y|do(A)]$

Motivation

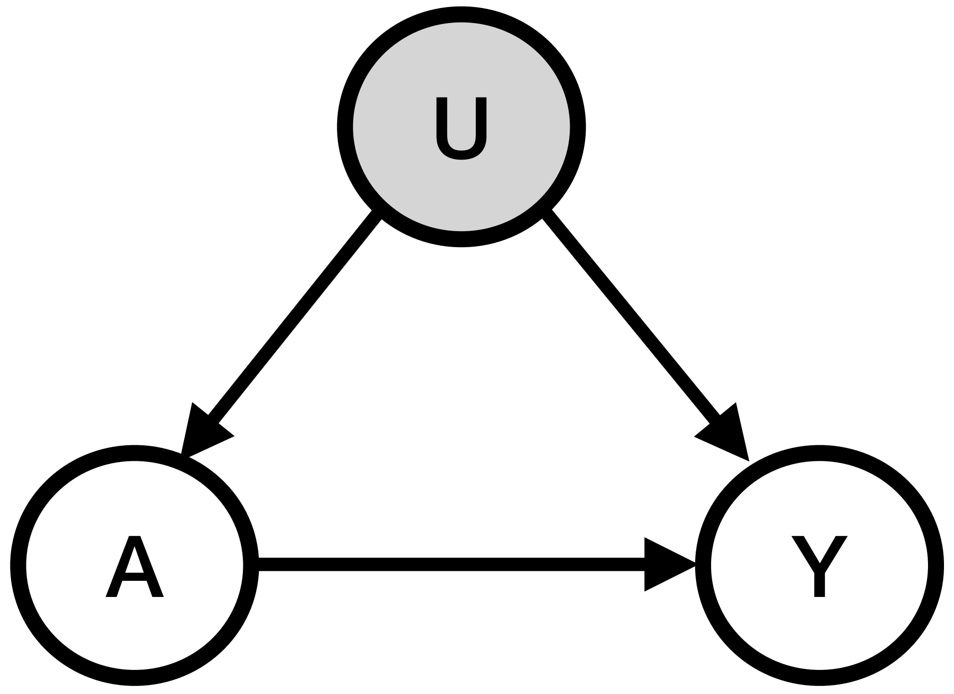
- Interesting causal questions often relates to notions imprecisely measured.
  - E.g. 1) fat intake on breast cancer, 2) skill mastery on career prospect.
- Two of the obstacles:
  - Measurement error on treatment.
  - Unobserved confounders.
- What if we don't account for them?



Want:
$\mathbb{E}[Y \mid do(A)]$

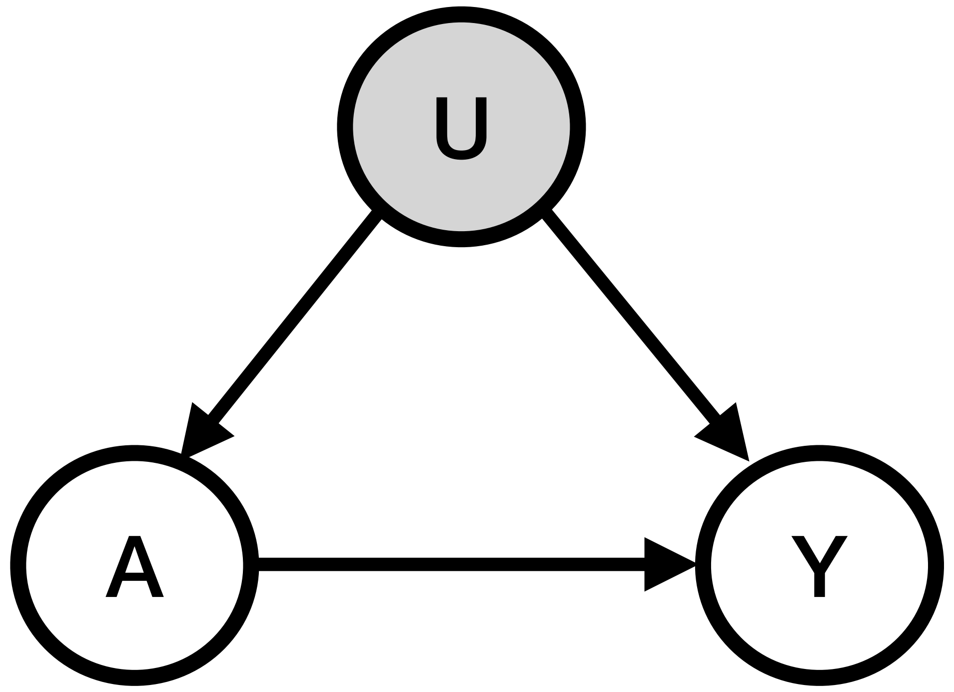
Consequences of confounding/measurement error?

Unobserved confounders:

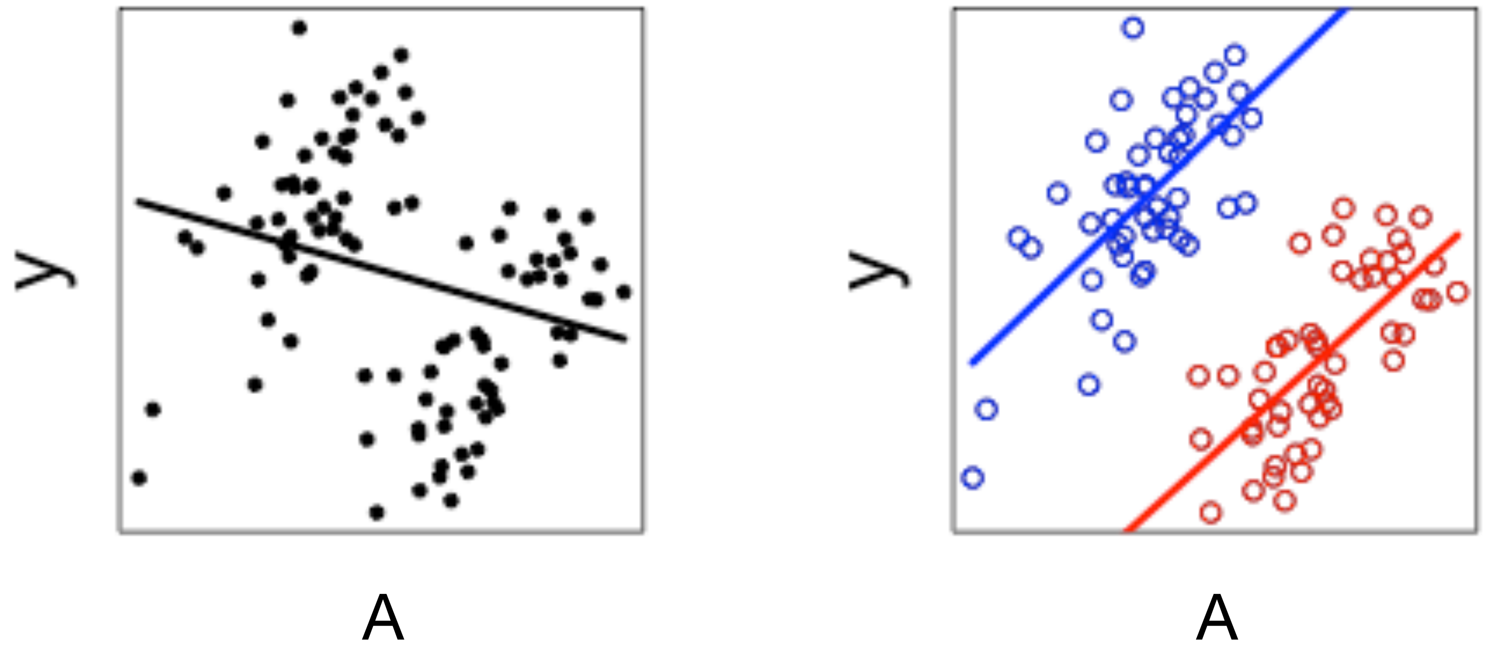


Consequences of confounding/measurement error?

Unobserved confounders:

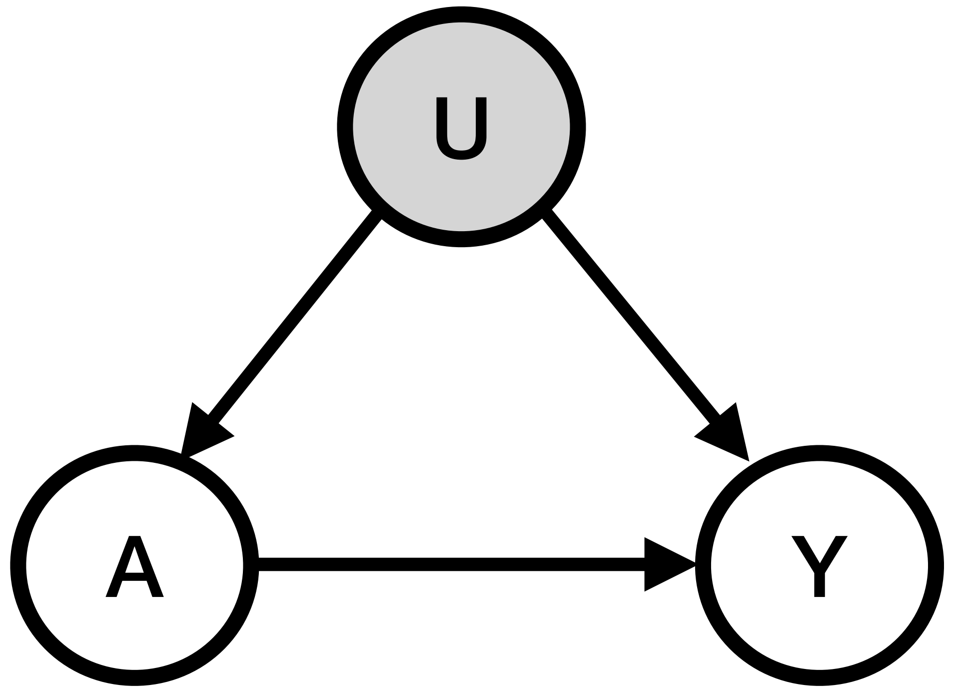


Simpson's paradox:

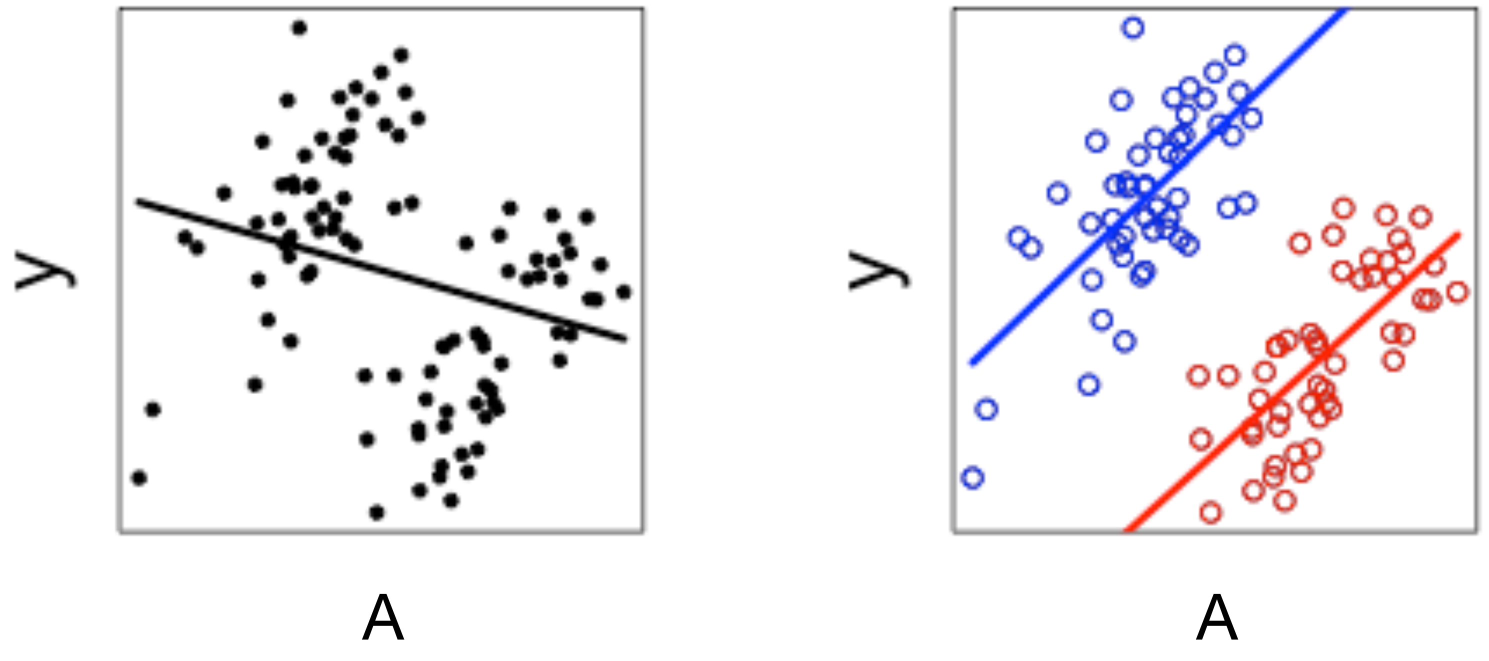


Consequences of confounding/measurement error?

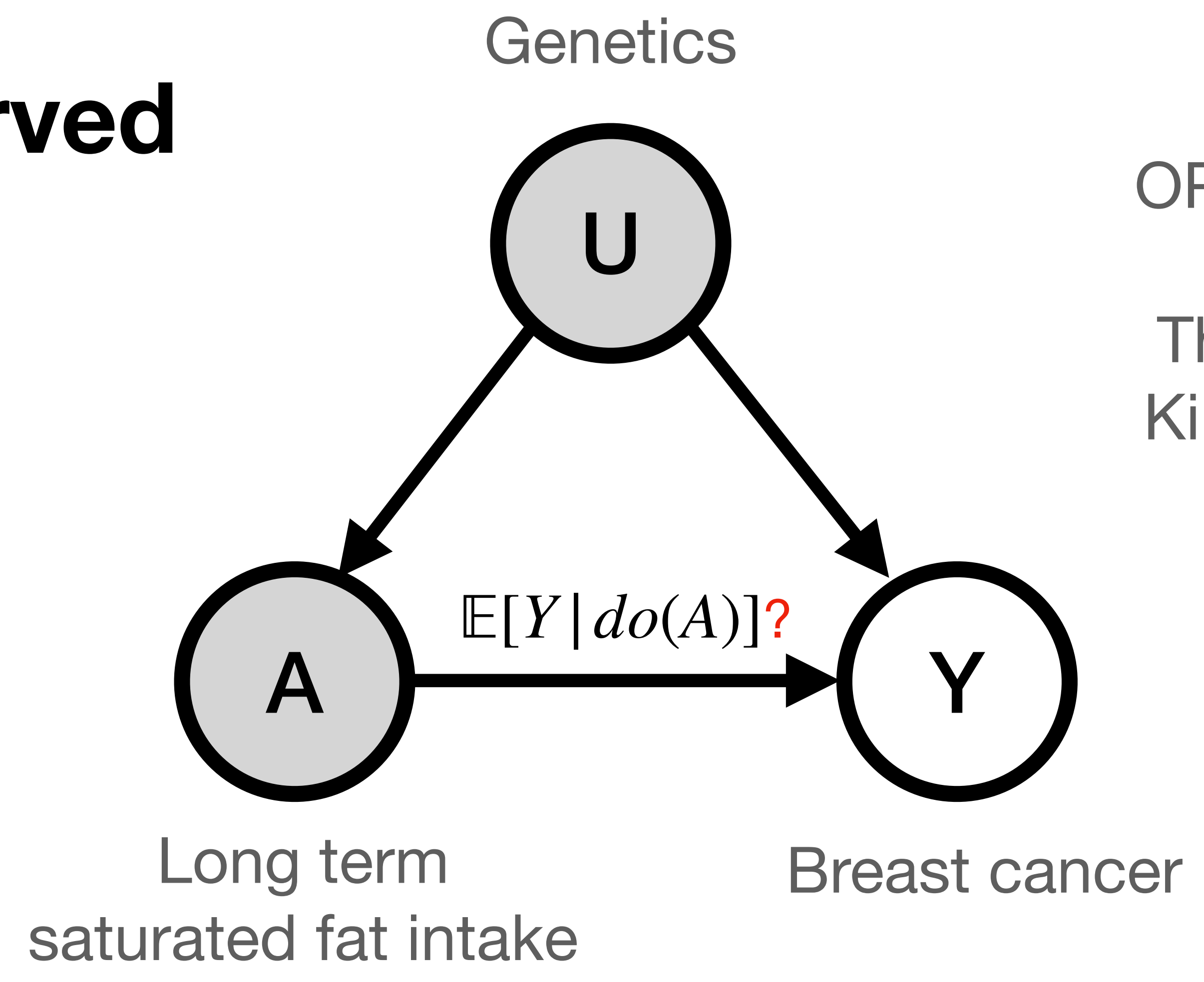
Unobserved confounders:



Simpson's paradox:



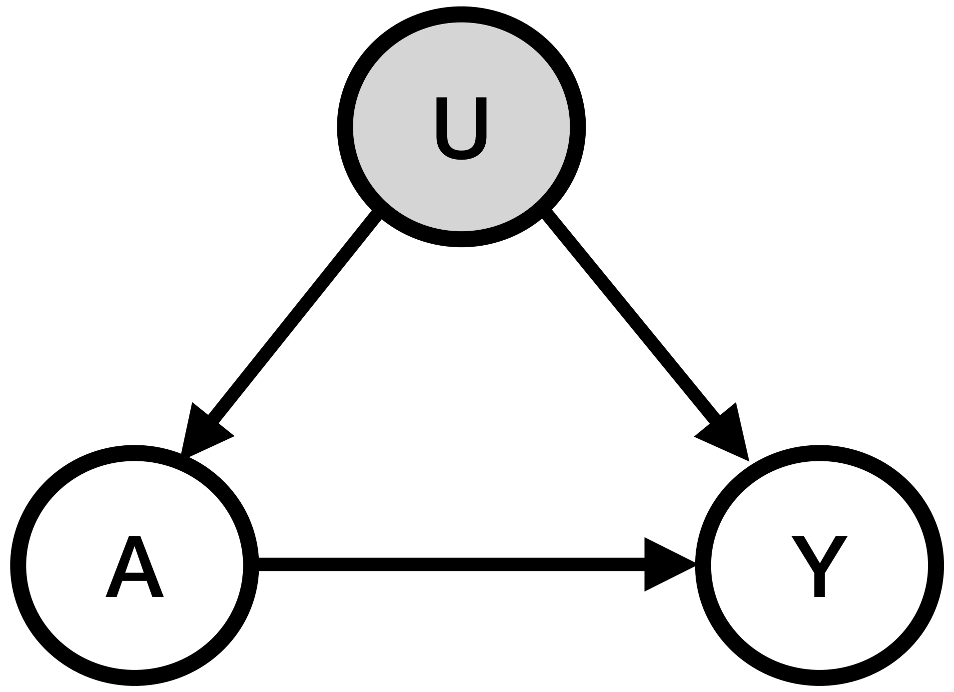
Action observed with error:



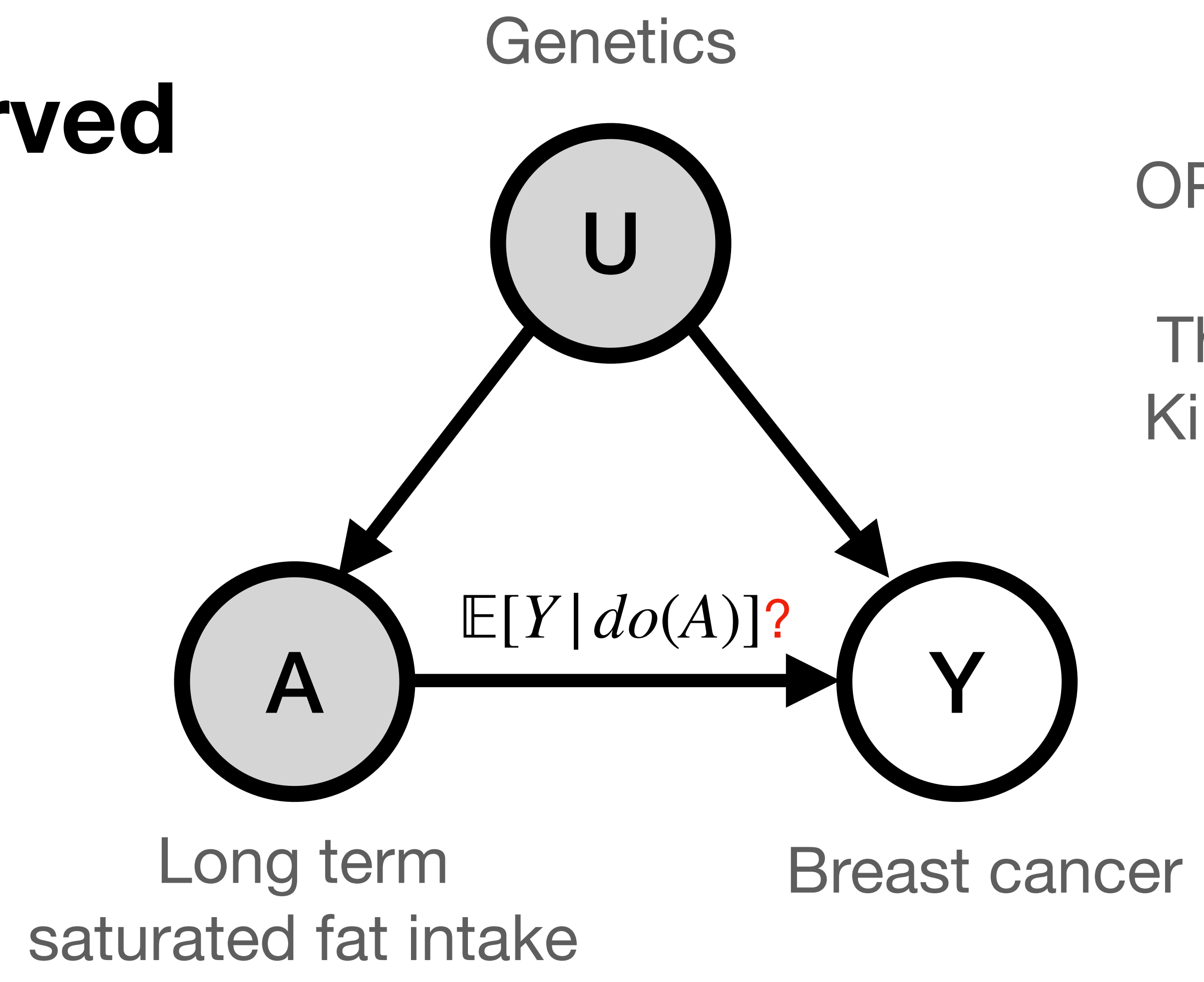
OPEN study: Subar, Thompson, Kipnis, et al. 2001

Consequences of confounding/measurement error?

Unobserved confounders:

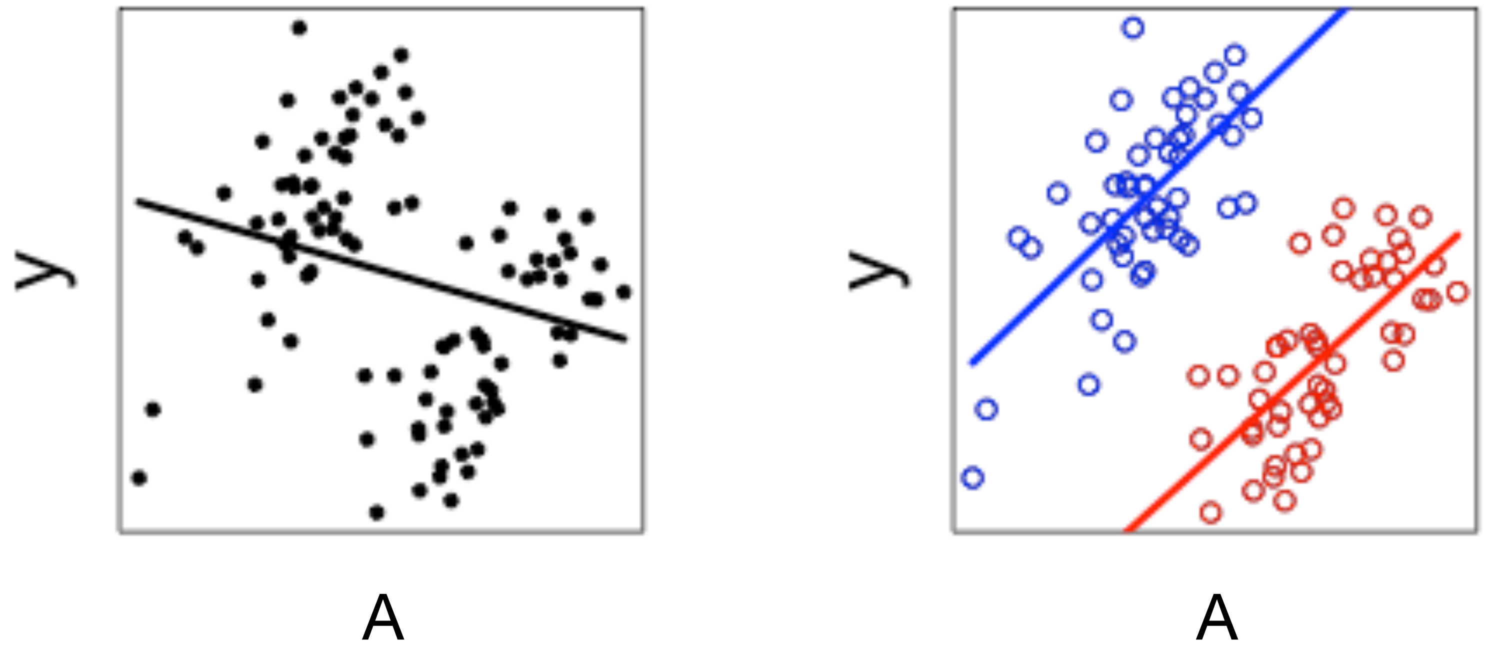


Action observed with error:

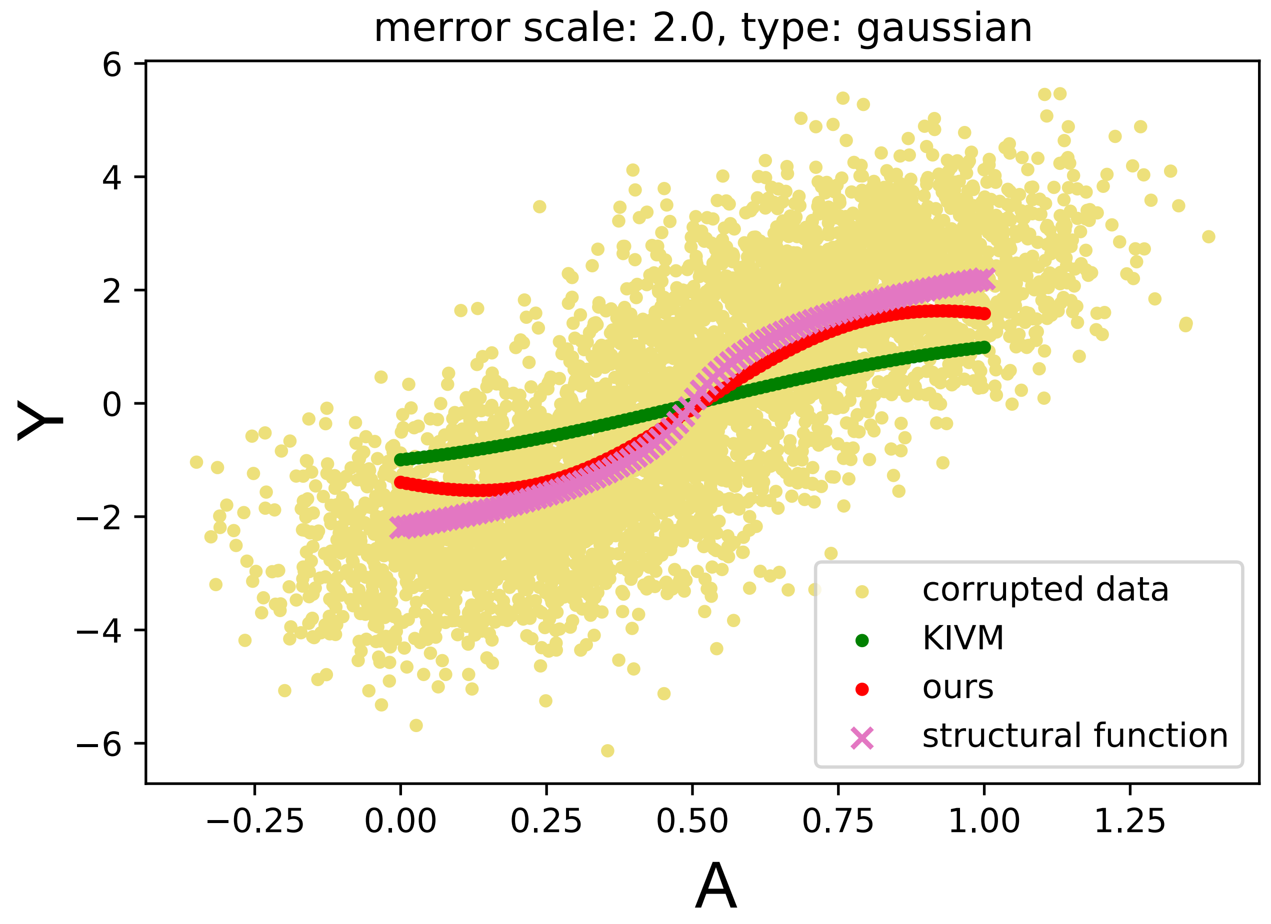


OPEN study: Subar, Thompson, Kipnis, et al. 2001

Simpson's paradox:

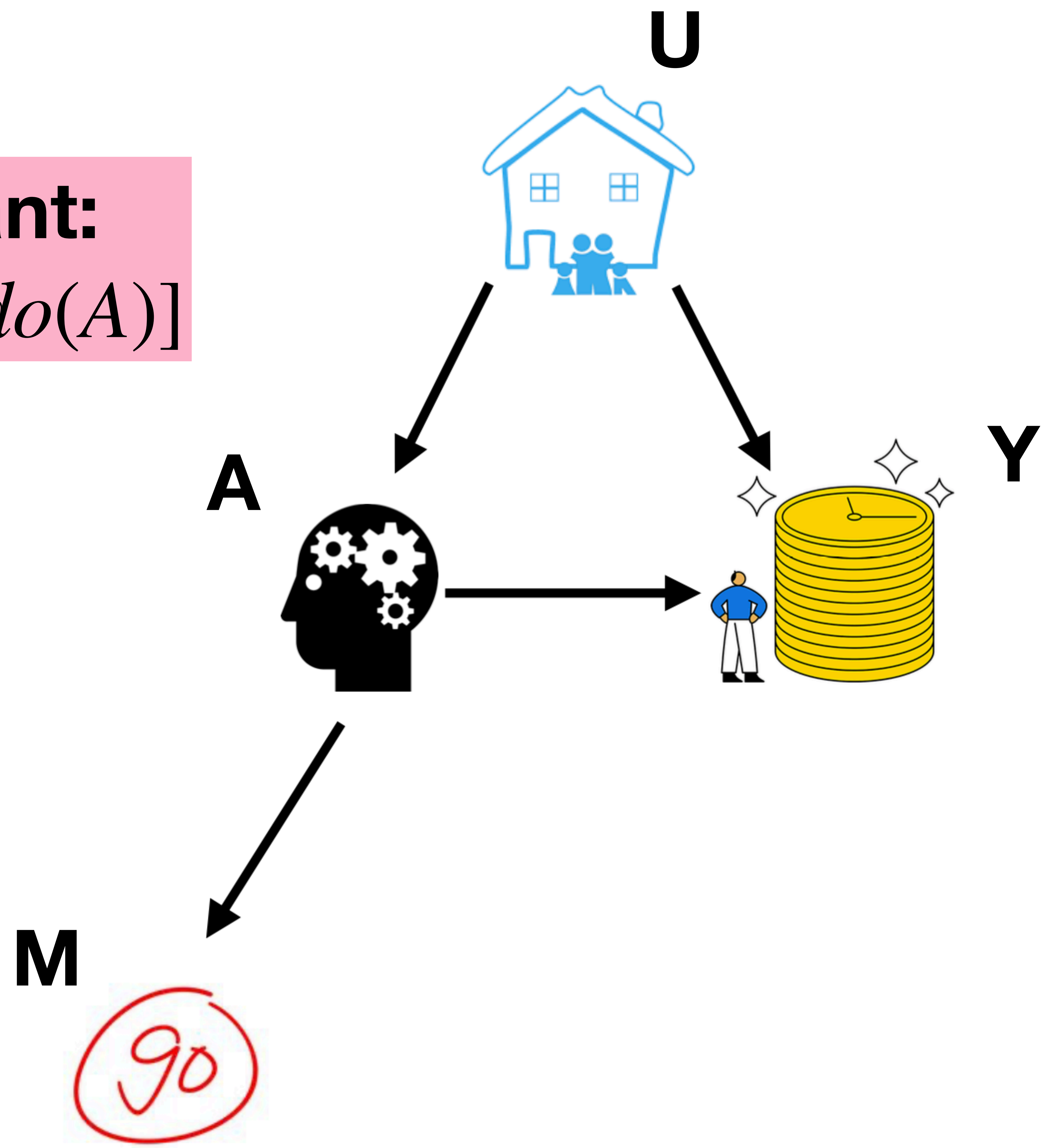


Mask interesting relationships:



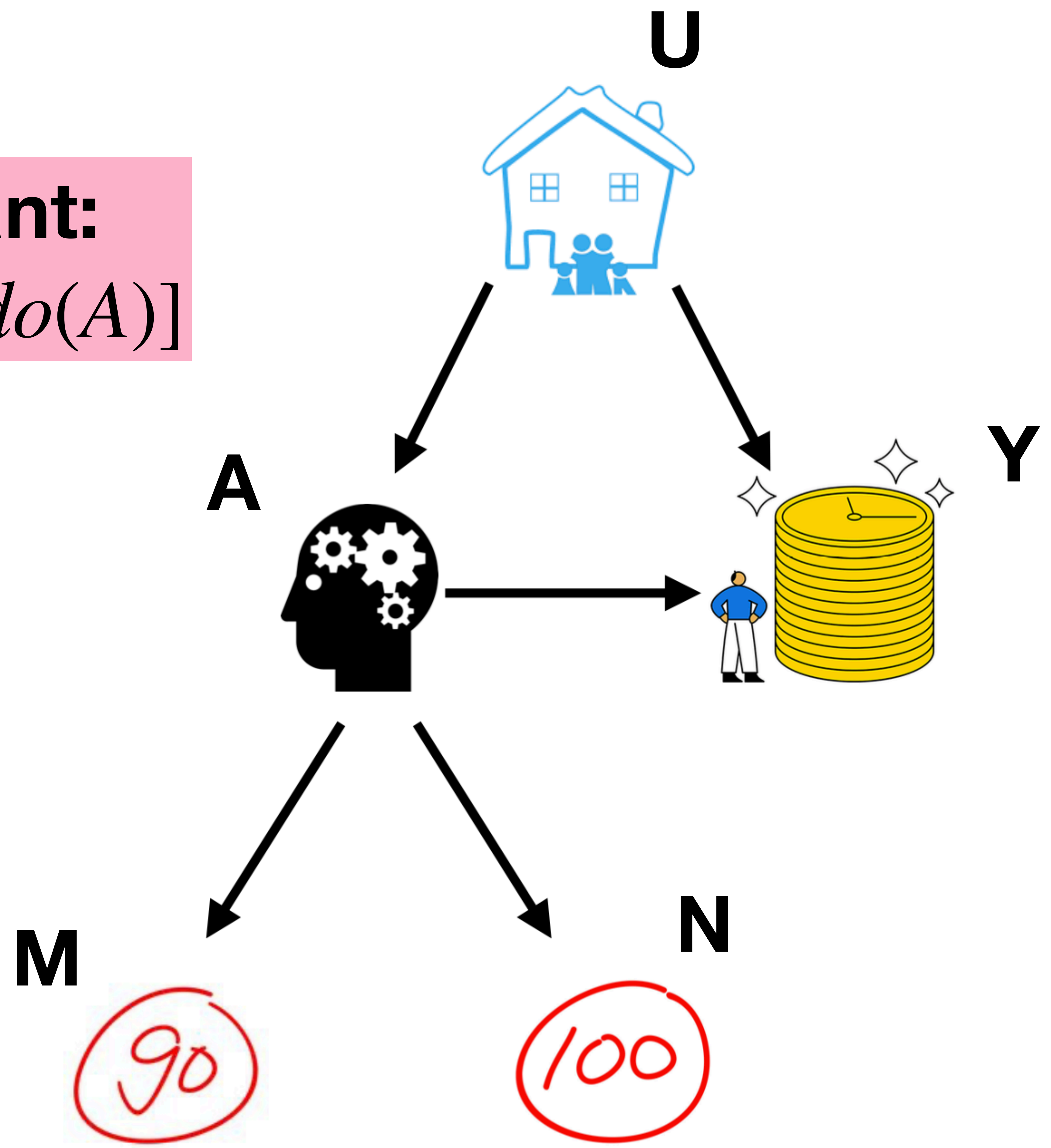
Measurement error on action variables - overview

Want:
$\mathbb{E}[Y \mid do(A)]$



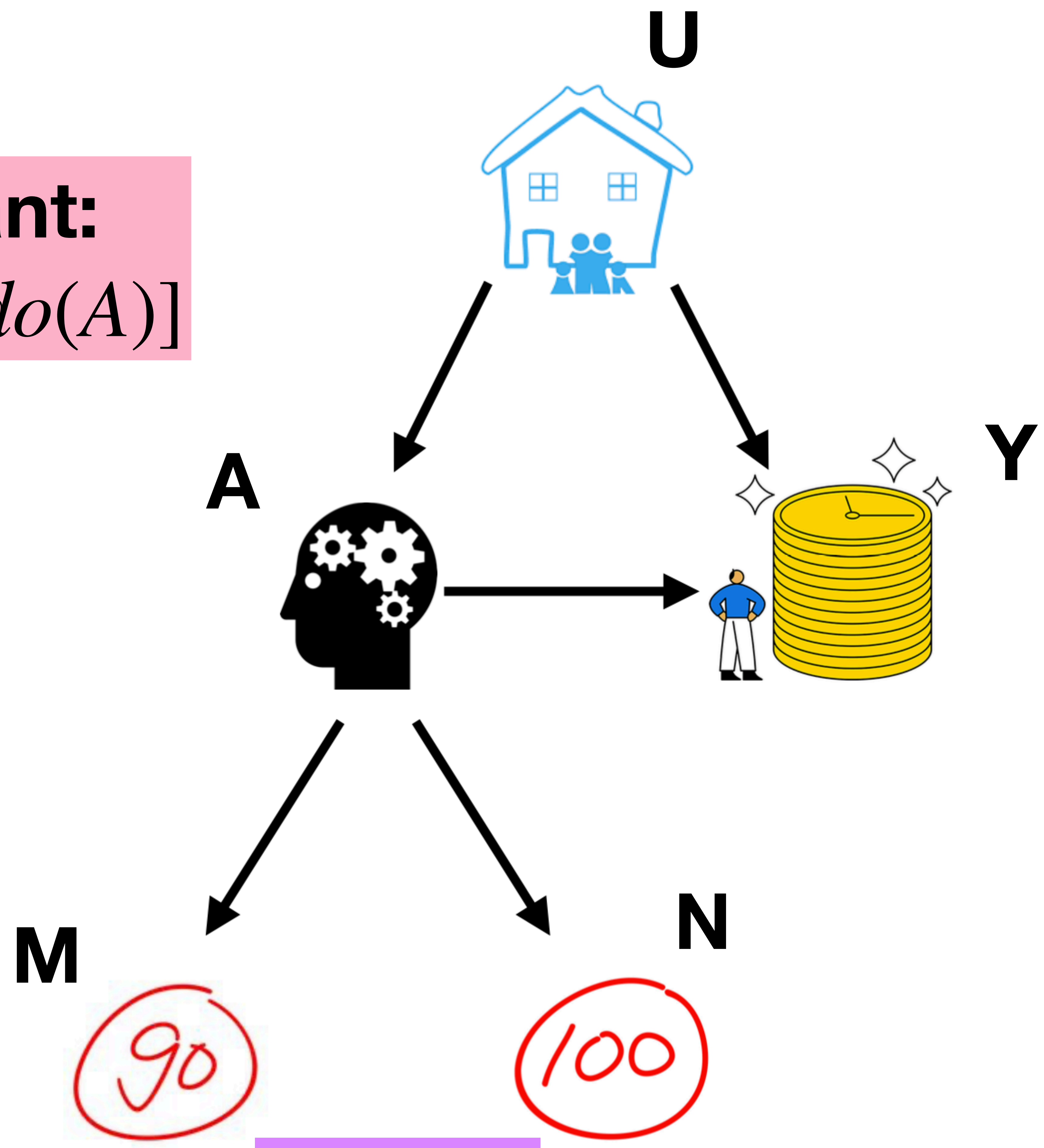
Measurement error on action variables - overview

Want:
$\mathbb{E}[Y \mid do(A)]$



Measurement error on action variables - overview

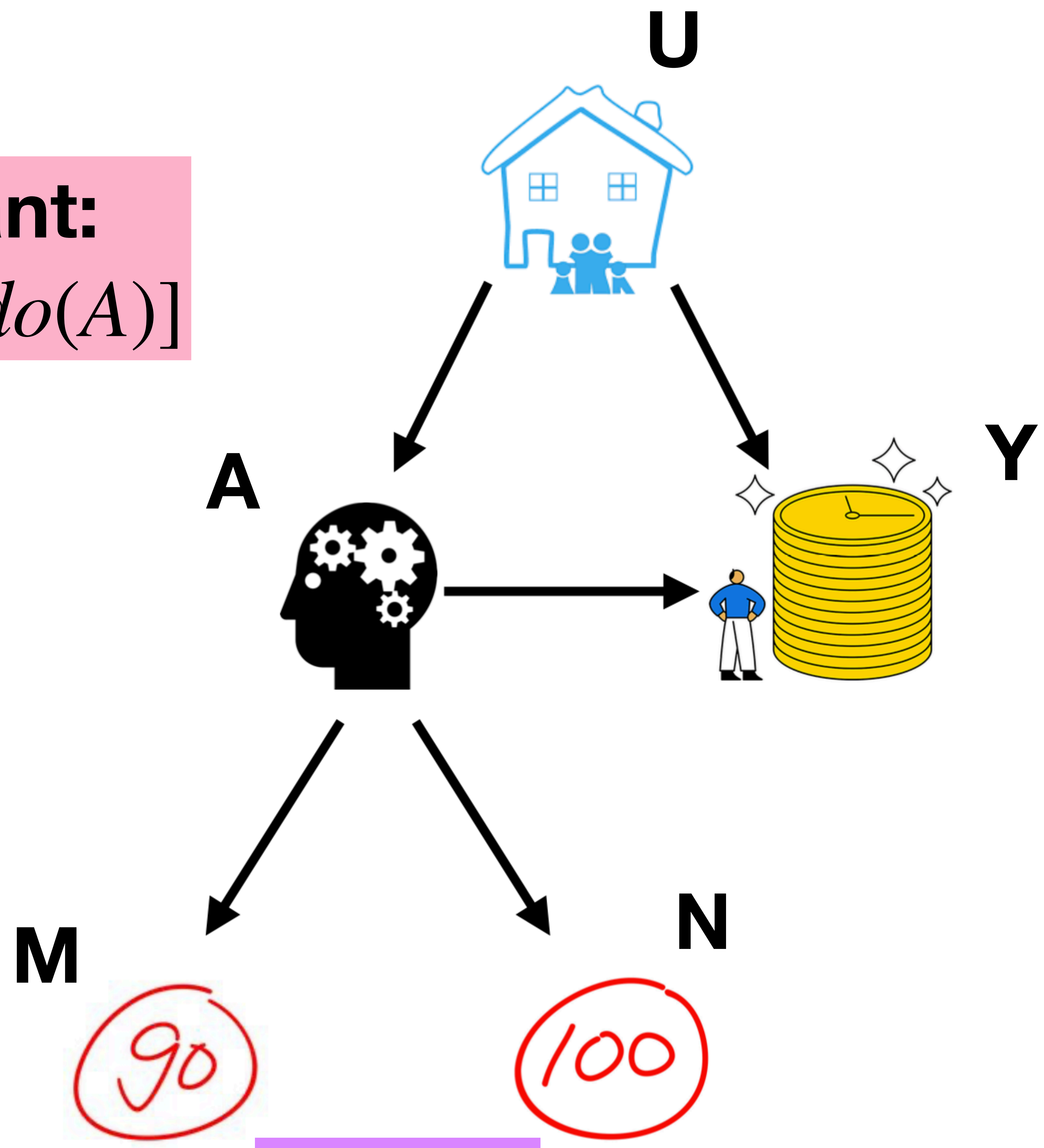
Want:
$\mathbb{E}[Y \mid do(A)]$



$M = A + \Delta M$
$N = A + \Delta N$

Measurement error on action variables - overview

Want:
$\mathbb{E}[Y \,|\, do(A)]$

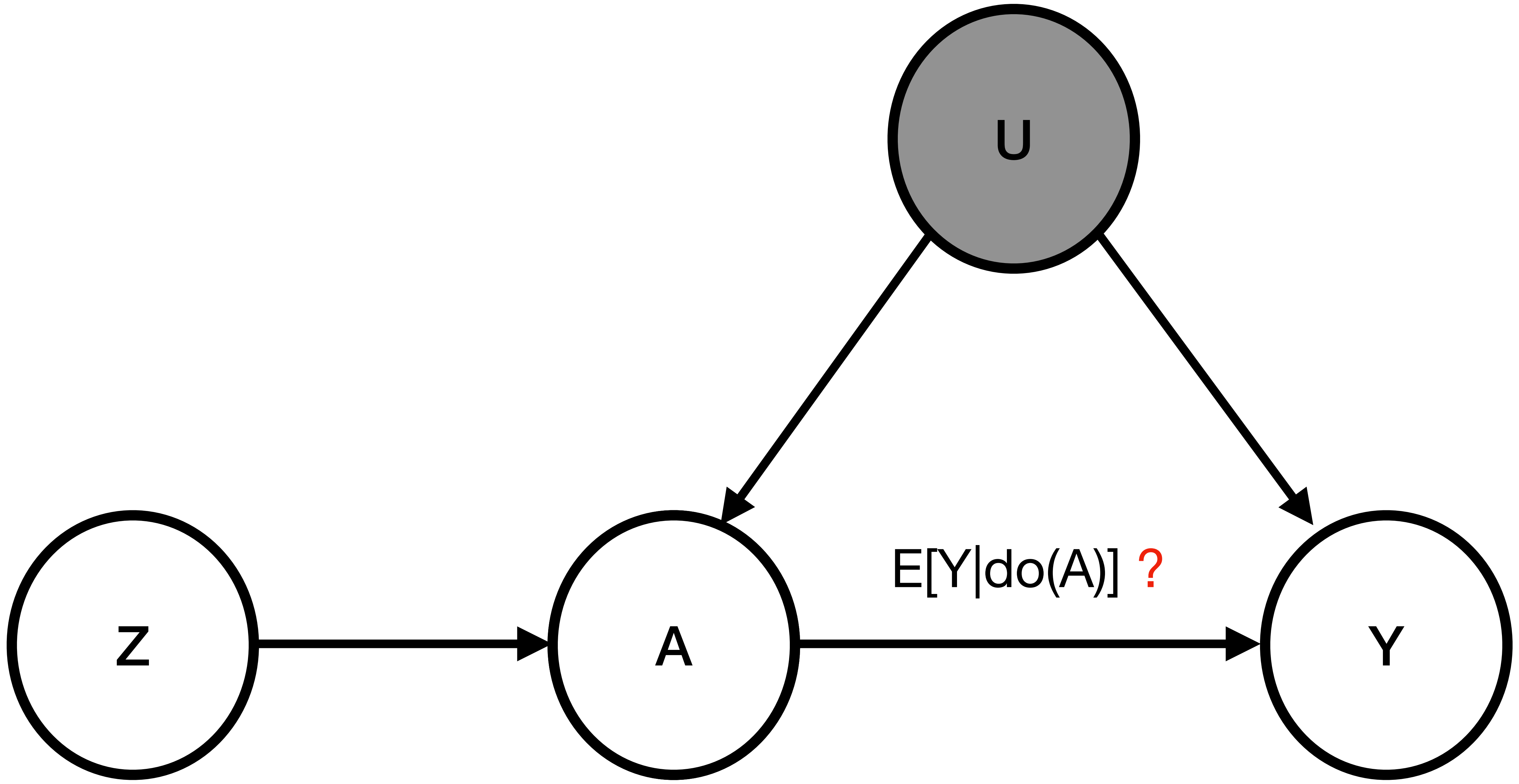


$M = A + \Delta M$
$N = A + \Delta N$

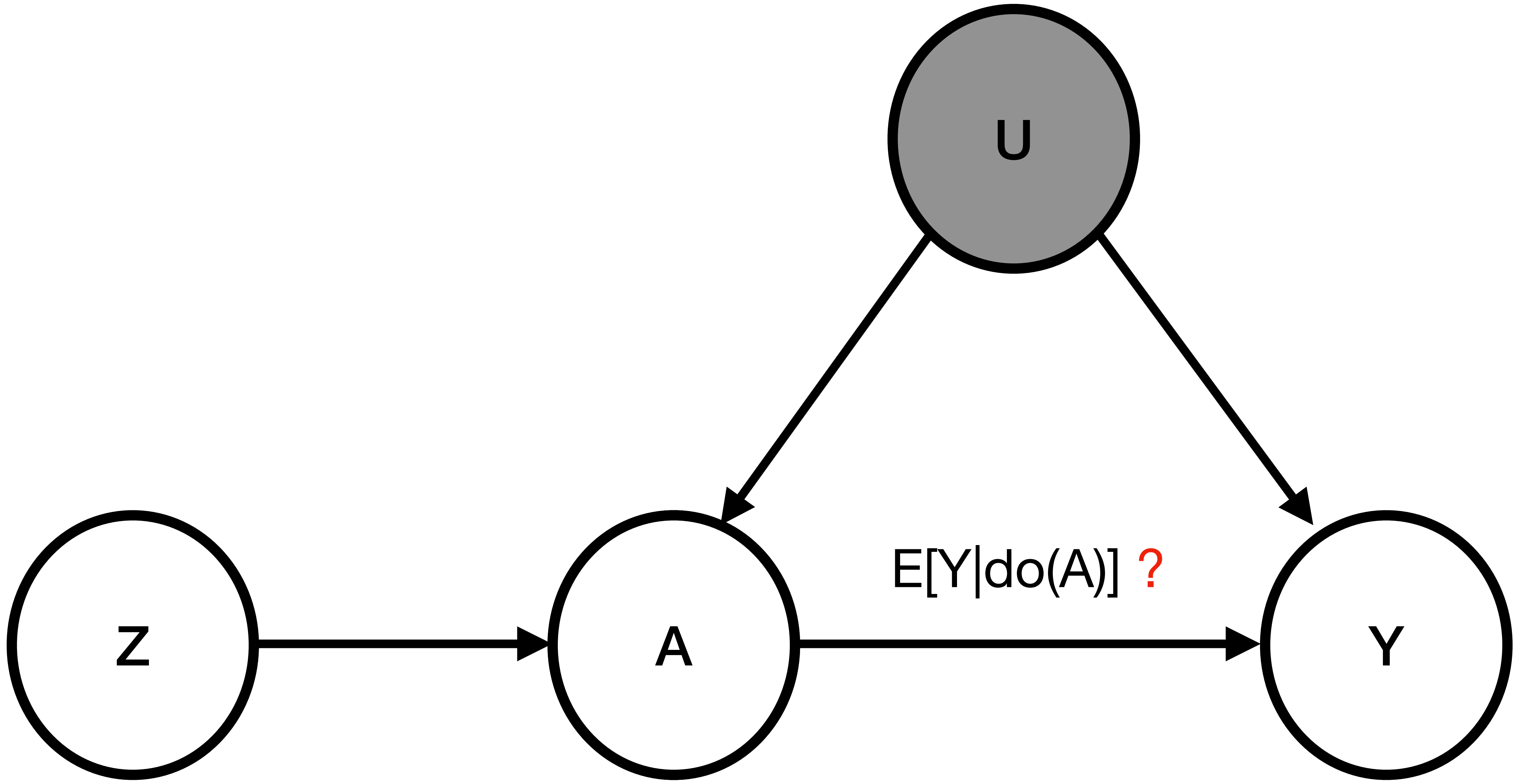
Kotlarski's:

$$\overbrace{\mathbb{E}_{\mathscr{P}_A}[e^{i\alpha A}]}^{\psi_A(\alpha)} = \exp\left(\int_0^\alpha i \frac{\mathbb{E}[Me^{i\nu N}]}{\mathbb{E}[e^{i\nu N}]} d\nu \right)$$

Identification with instrumental variables

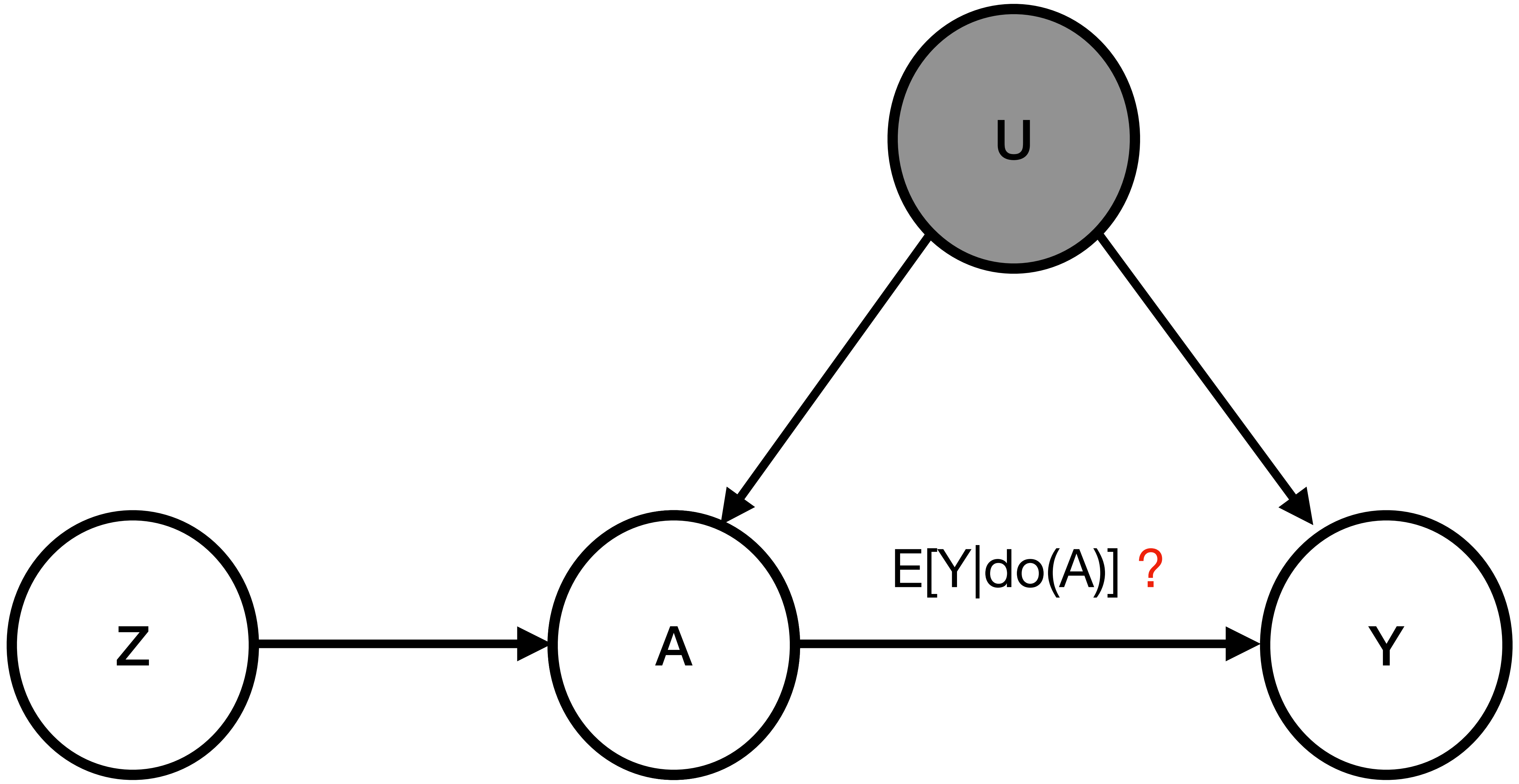


Identification with instrumental variables



Identification:

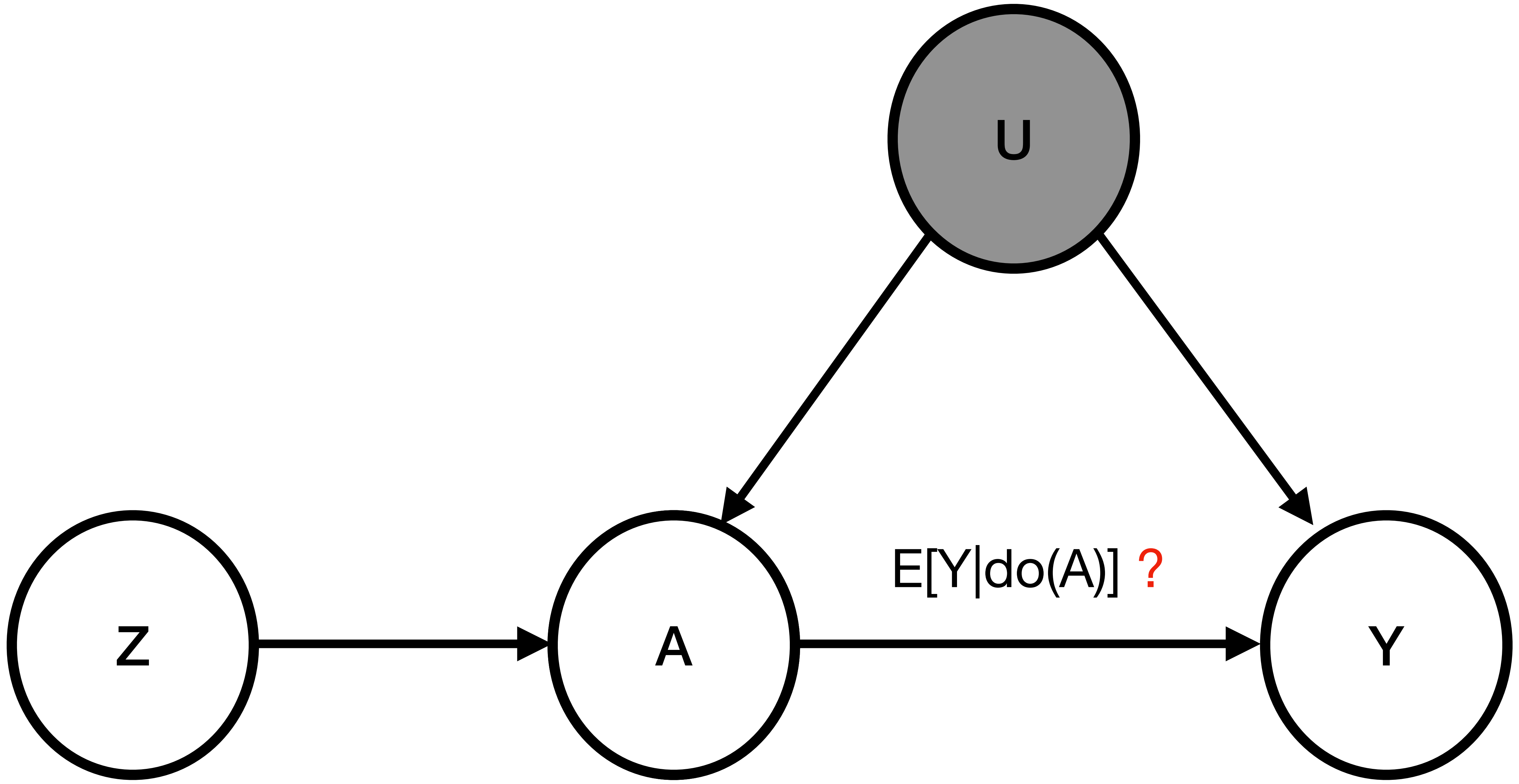
Identification with instrumental variables



Identification:

$$Y = f(A) + U \quad \mathbb{E}[U \mid Z] = 0$$

Identification with instrumental variables

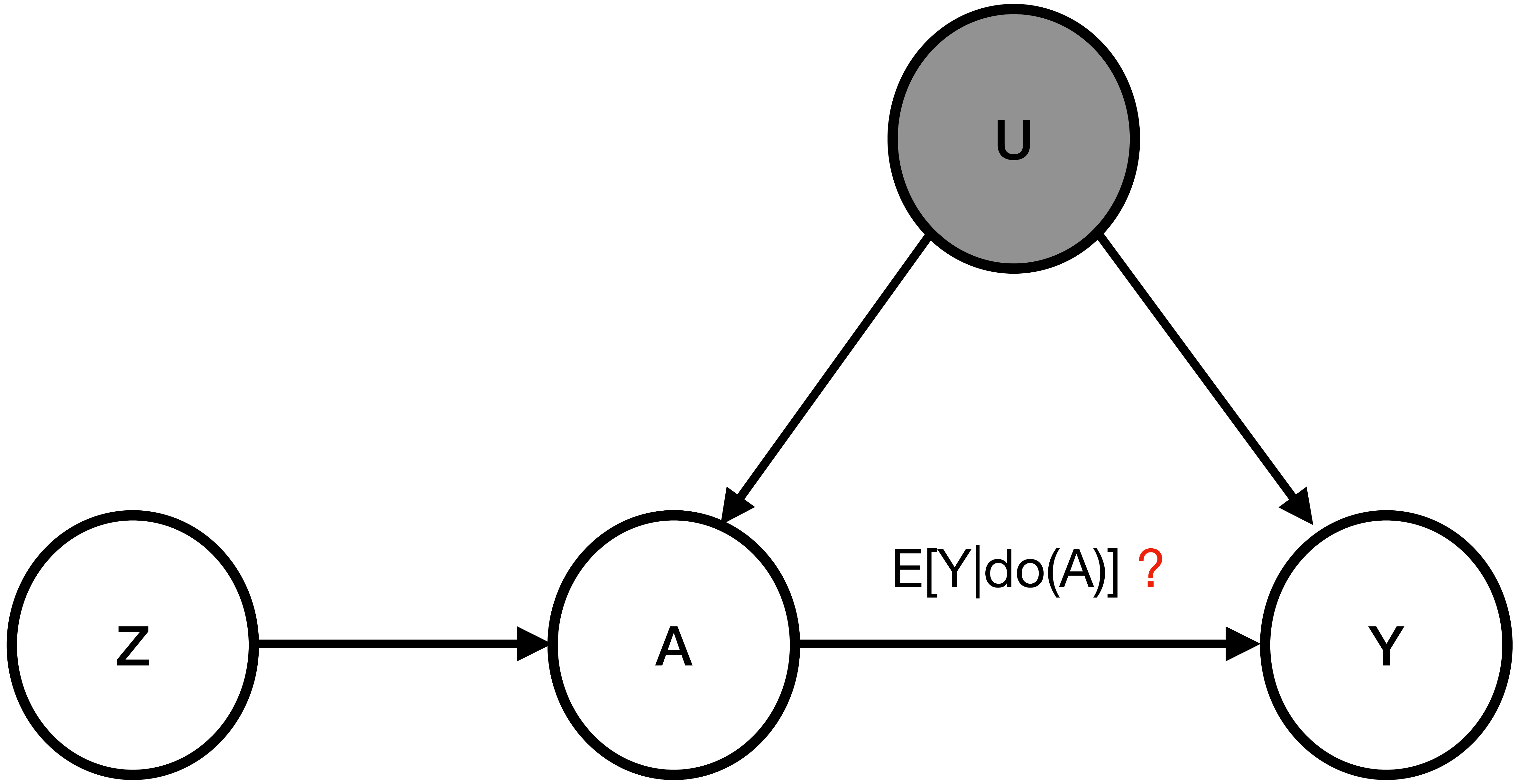


Identification:

$$Y = f(A) + U \quad \mathbb{E}[U \mid Z] = 0$$

$$f(A) = \mathbb{E}[Y \mid do(A)]$$

Identification with instrumental variables



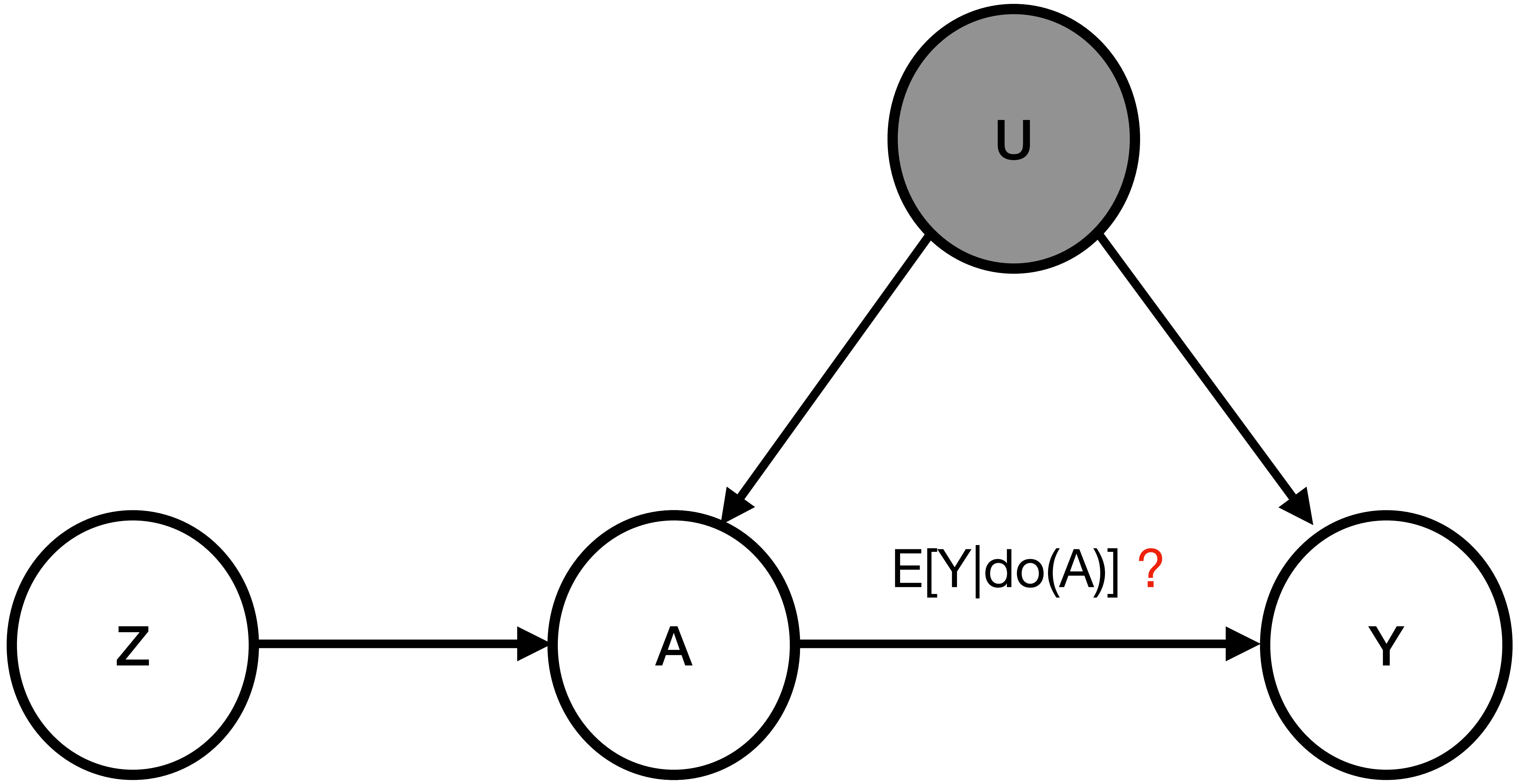
Identification:

$$Y = f(A) + U \quad \mathbb{E}[U \mid Z] = 0$$

$$f(A) = \mathbb{E}[Y \mid do(A)]$$

$$\mathbb{E}[Y \mid Z] = \int_{\mathscr{A}} f(a)p(a \mid Z)da$$

Identification with instrumental variables



Identification:

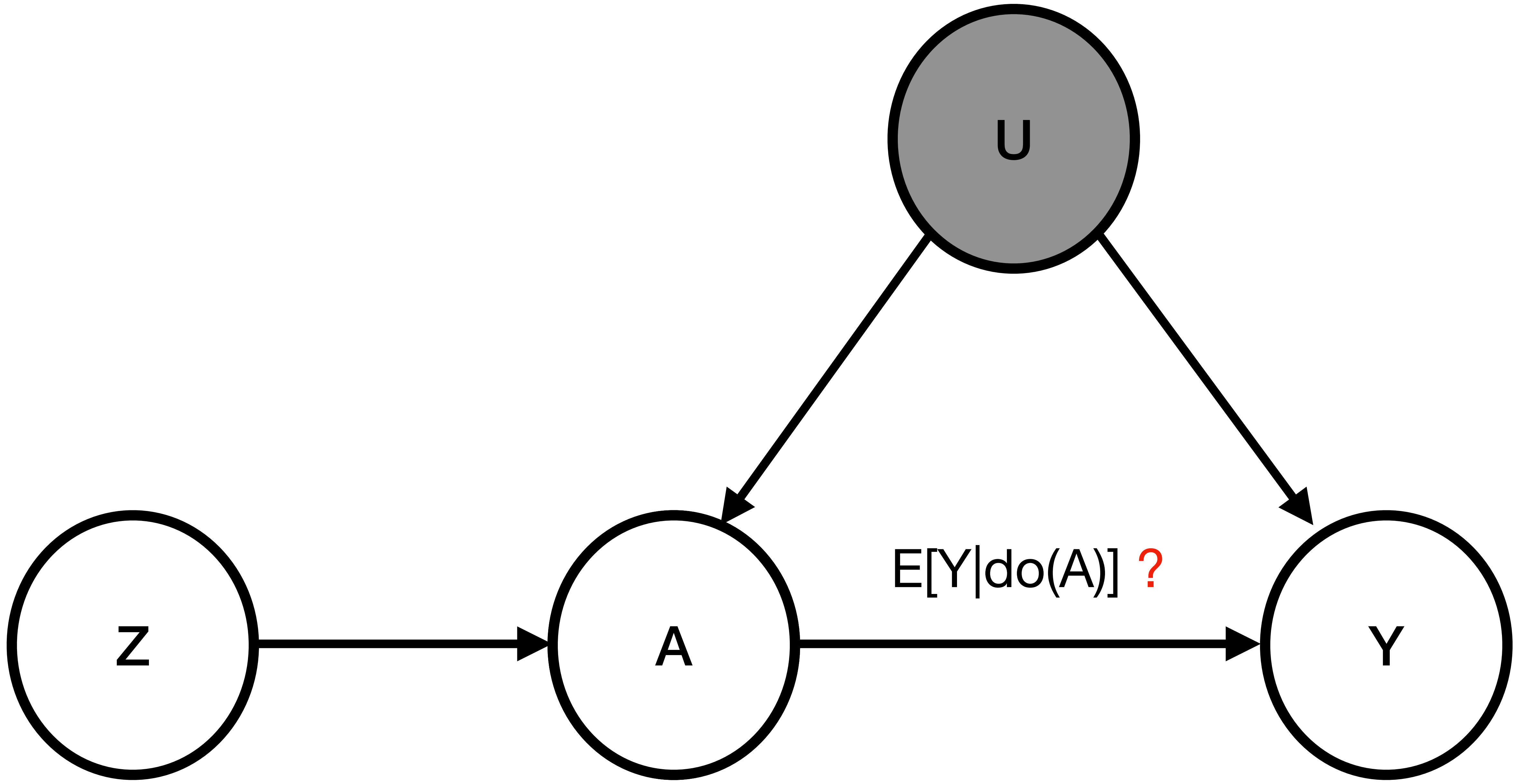
$$Y = f(A) + U \quad \mathbb{E}[U \mid Z] = 0$$

$$f(A) = \mathbb{E}[Y \mid do(A)]$$

$$\mathbb{E}[Y \mid Z] = \int_{\mathcal{A}} f(a)p(a \mid Z)da$$

???

Identification with instrumental variables



Identification:

$$Y = f(A) + U \quad \mathbb{E}[U | Z] = 0$$

$$f(A) = \mathbb{E}[Y | do(A)]$$

$$\mathbb{E}[Y | Z] = \int_{\mathscr{A}} f(a)p(a | Z)da$$

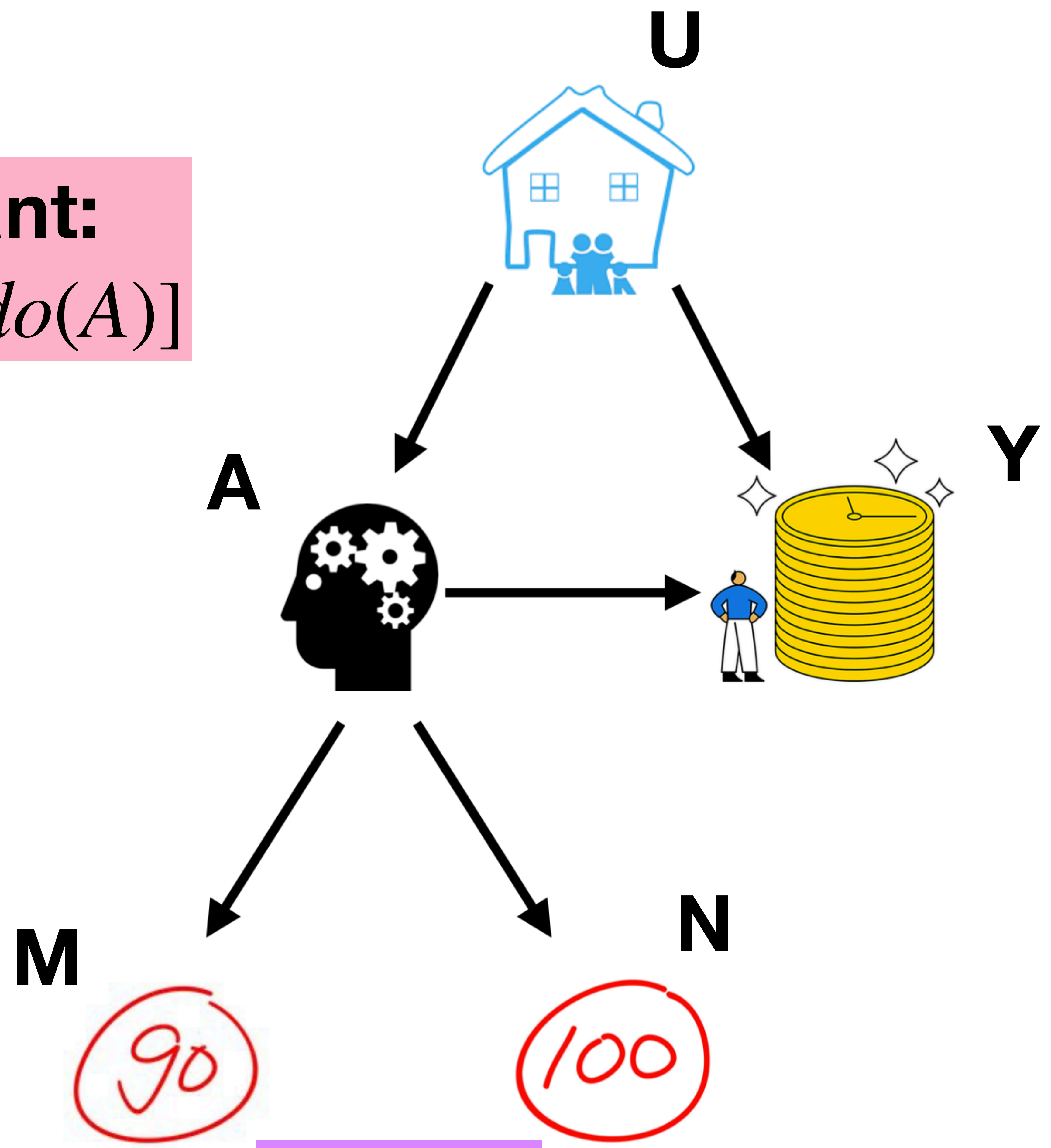
???

But if $f(a) = \theta^T \phi(a)$, then rhs simplies to

$$\mathbb{E}[Y | Z] = \theta^T \mathbb{E}[\phi(A) | Z]$$

Measurement error on action variables - overview

Want:
$\mathbb{E}[Y \mid do(A)]$



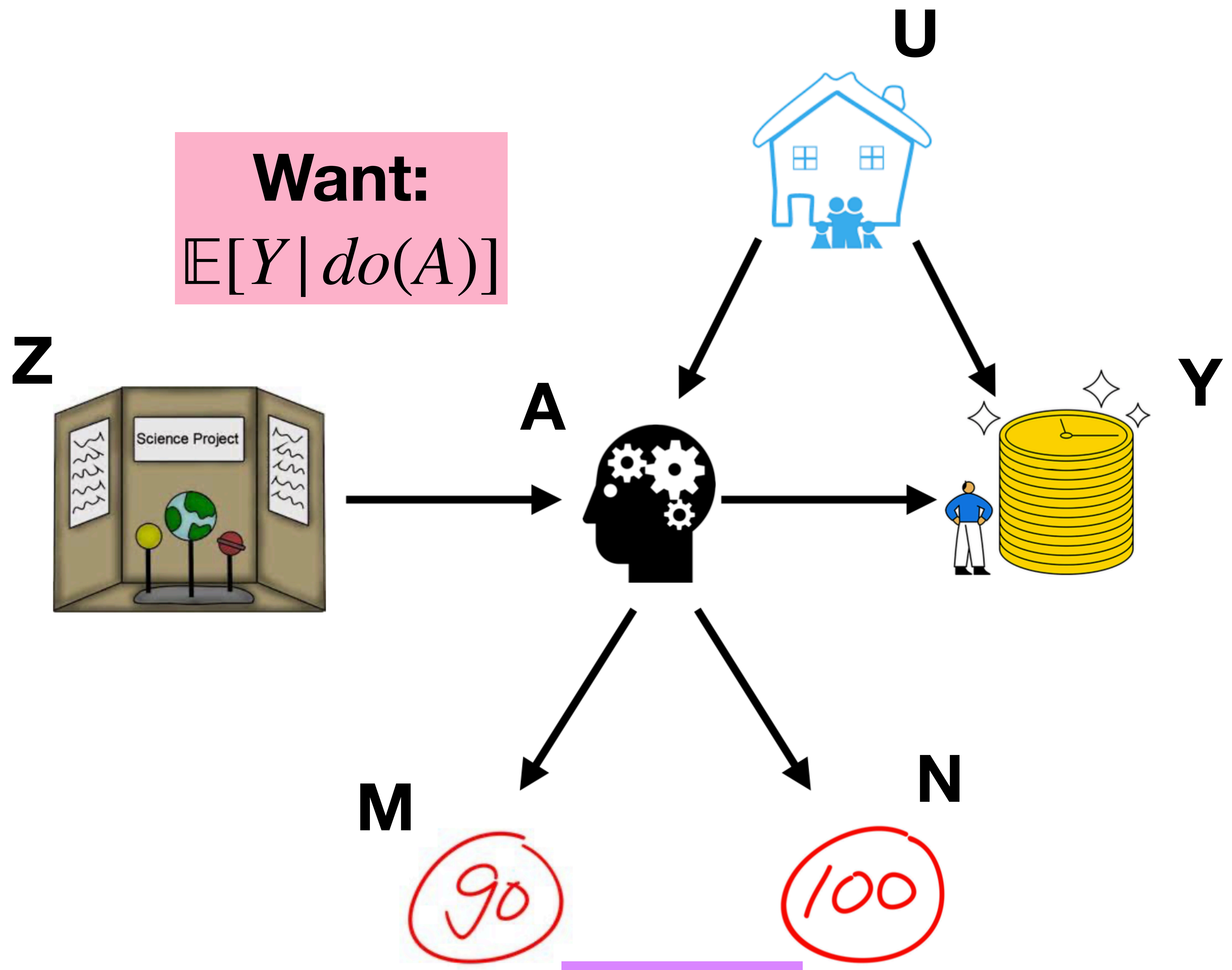
$M = A + \Delta M$
$N = A + \Delta N$

Kotlarski's:

$$\overbrace{\mathbb{E}_{\mathscr{P}_A}[e^{i\alpha A}]}^{\psi_A(\alpha)} = \exp\left(\int_0^\alpha i\frac{\mathbb{E}[Me^{i\nu N}]}{\mathbb{E}[e^{i\nu N}]}d\nu\right)$$

Measurement error on action variables - overview

Want:
$\mathbb{E}[Y \mid do(A)]$



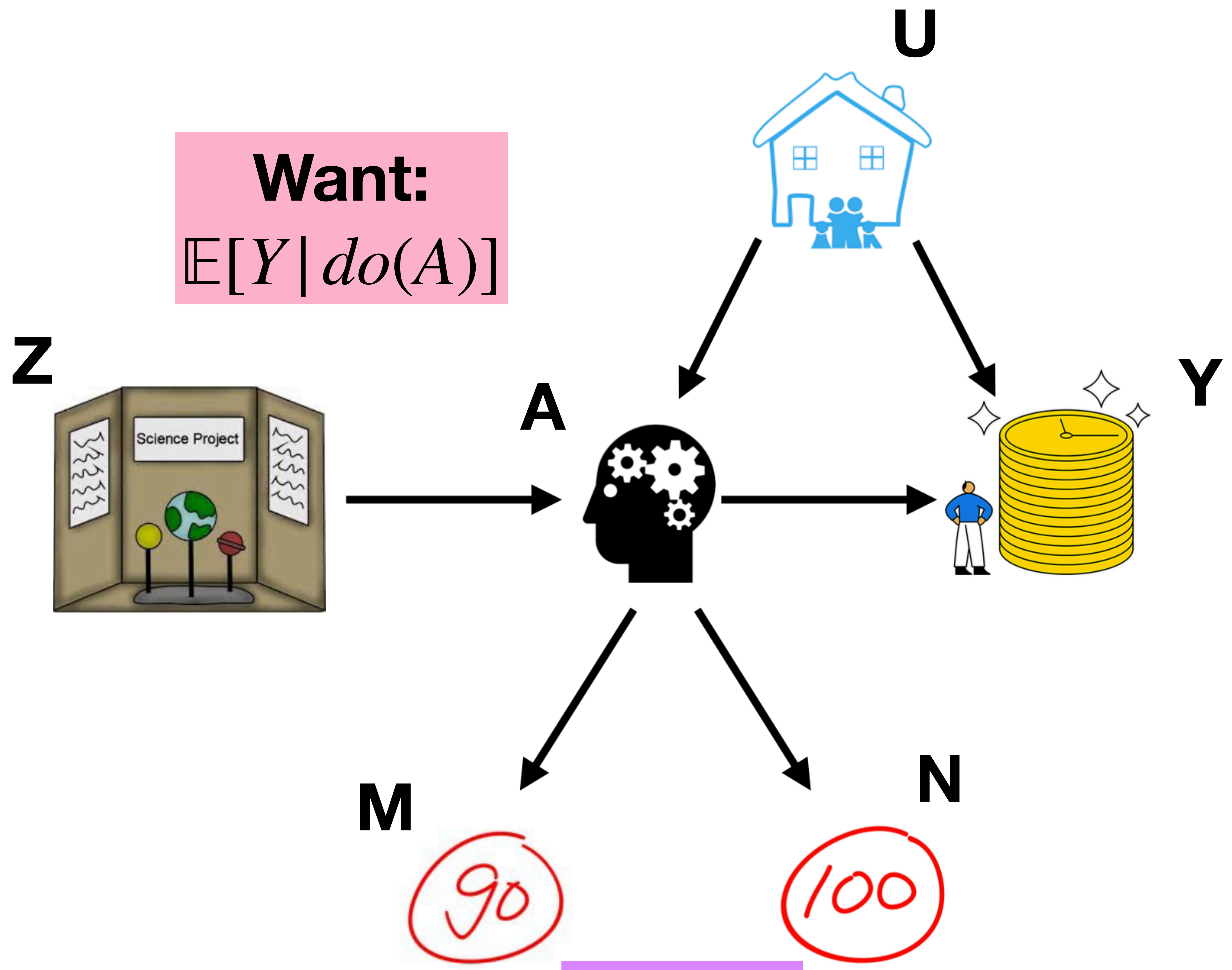
$M = A + \Delta M$
$N = A + \Delta N$

Kotlarski's:

$$\overbrace{\mathbb{E}[e^{i\alpha A} \mid z]}^{\psi_{A|z}(\alpha)} = \exp\left(\int_0^{\alpha} i \frac{\mathbb{E}[M e^{i\nu N} \mid z]}{\mathbb{E}[e^{i\nu N} \mid z]} d\nu \right)$$

Measurement error on action variables - overview

Want:
$\mathbb{E}[Y \mid do(A)]$



$M = A + \Delta M$
$N = A + \Delta N$

Kotlarski's**:**

$$\overbrace{\mathbb{E}[e^{i\alpha A} \mid z]}^{\psi_{A|z}(\alpha)} = \exp\left(\int_0^\alpha i \frac{\mathbb{E}[Me^{i\nu N} \mid z]}{\mathbb{E}[e^{i\nu N} \mid z]} d\nu \right)$$

What about
$\mu_{A|z} := \mathbb{E}[\phi(A) \mid z]$**?**

From $\hat{\psi}^n_{X|z}(\alpha)$ to $\hat{\mu}^n_{X|z}(y)$ (= $\mathbb{E}[\phi(X) \,|\, z]$)

From $\hat{\psi}^n_{X|z}(\alpha)$ to $\hat{\mu}^n_{X|z}(y)$ (= $\mathbb{E}[\phi(X) \,|\, z]$)

Have $\hat{\mu}^n_{X|z}(y) = \sum_{j=1}^{n} \hat{\gamma}^n_j(z) k(x_j, y)$.

From $\hat{\psi}^n_{X|z}(\alpha)$ to $\hat{\mu}^n_{X|z}(y)$ (= $\mathbb{E}[\phi(X) \mid z]$)

Have $\hat{\mu}^n_{X|z}(y) = \sum_{j=1}^{n} \hat{\gamma}^n_j(z) k(x_j, y)$.

Where $\hat{\gamma}^n_j(z) = (K_{ZZ} + n\hat{\lambda}^n I)^{-1} K_{Zz}$.

From $\hat{\psi}^n_{X|z}(\alpha)$ to $\hat{\mu}^n_{X|z}(y)$ ($= \mathbb{E}[\phi(X) \mid z]$)

Have $\hat{\mu}^n_{X|z}(y) = \sum_{j=1}^{n} \hat{\gamma}^n_j(z) k(x_j, y)$.

Where $\hat{\gamma}^n_j(z) = (K_{ZZ} + n\hat{\lambda}^n I)^{-1} K_{Zz}$.

Let $\hat{\psi}^n_{X|z}(\alpha) := \sum_{j=1}^{n} \hat{\gamma}^n_j(z) e^{i\alpha x_j}$.

From $\hat{\psi}^n_{X|z}(\alpha)$ to $\hat{\mu}^n_{X|z}(y)$ (= $\mathbb{E}[\phi(X)\,|\,z]$)

Have $\hat{\mu}^n_{X|z}(y) = \sum_{j=1}^{n} \hat{\gamma}^n_j(z) k(x_j, y)$.

Where $\hat{\gamma}^n_j(z) = (K_{ZZ} + n\hat{\lambda}^n I)^{-1} K_{Zz}$.

Let $\hat{\psi}^n_{X|z}(\alpha) := \sum_{j=1}^{n} \hat{\gamma}^n_j(z) e^{i\alpha x_j}$.

Theorem 1. With translation-invariant, characteristic kernel:

$\hat{\mu}^n_{X|Z} \to^n \mu_{X|Z}$ iff $\hat{\psi}^n_{X|Z} \to^n \psi_{X|Z}$ in IFT of kernel.

From $\hat{\psi}^n_{X|z}(\alpha)$ to $\hat{\mu}^n_{X|z}(y)$ (= $\mathbb{E}[\phi(X) \mid z]$)

Have $\hat{\mu}^n_{X|z}(y) = \sum_{j=1}^{n} \hat{\gamma}^n_j(z) k(x_j, y)$.

Where $\hat{\gamma}^n_j(z) = (K_{ZZ} + n\hat{\lambda}^n I)^{-1} K_{Zz}$.

Let $\hat{\psi}^n_{X|z}(\alpha) := \sum_{j=1}^{n} \hat{\gamma}^n_j(z) e^{i\alpha x_j}$.

Theorem 1. With translation-invariant, characteristic kernel:

$\hat{\mu}^n_{X|Z} \to^n \mu_{X|Z}$ iff $\hat{\psi}^n_{X|Z} \to^n \psi_{X|Z}$ in IFT of kernel.

$$\overbrace{\mathbb{E}_{\mathscr{P}_{A|z}}[e^{i\alpha X}](\alpha)}^{\psi_{A|z}(\alpha)} = \exp\left(\int_0^\alpha i \frac{\overbrace{\mathbb{E}[M e^{i\nu N} \mid z]}^{\frac{\partial}{\partial \upsilon}\psi_{M,N|z}(\upsilon,\nu)\big|_{\upsilon=0}}}{\underbrace{\mathbb{E}[e^{i\nu N} \mid z]}_{\psi_{N|z}(\nu)}} d\nu \right)$$

Measurement Error KIV

To obtain $\hat{\psi}^n_{A|z}$:

$$\overbrace{\mathbb{E}_{\mathscr{P}_{A|z}}[e^{i\alpha X}](\alpha)}^{\psi_{A|z}(\alpha)} = \exp\left(\int_0^\alpha i \frac{\overbrace{\mathbb{E}[Me^{i\nu N} \,|\, z]}^{\frac{\partial}{\partial \upsilon}\psi_{M,N|z}(\upsilon,\nu)\big|_{\upsilon=0}}}{\underbrace{\mathbb{E}[e^{i\nu N} \,|\, z]}_{\psi_{N|z}(\nu)}} d\nu \right) \qquad (1)$$

Measurement Error KIV

To obtain $\hat{\psi}^n_{A|z}$:

$$\overbrace{\mathbb{E}_{\mathscr{P}_{A|z}}[e^{i\alpha X}](\alpha)}^{\psi_{A|z}(\alpha)} = \exp\left(\int_0^\alpha i\frac{\overbrace{\mathbb{E}[Me^{i\nu N}\,|\,z]}^{\frac{\partial}{\partial v}\psi_{M,N|z}(v,\nu)\big|_{v=0}}}{\underbrace{\mathbb{E}[e^{i\nu N}\,|\,z]}_{\psi_{N|z}(\nu)}}d\nu\right) \qquad (1)$$

Differentiate wrt α to remove integral.

Measurement Error KIV

To obtain $\hat{\psi}^n_{A|z}$:

$$\overbrace{\mathbb{E}_{\mathscr{P}_{A|z}}[e^{i\alpha X}](\alpha)}^{\psi_{A|z}(\alpha)} = \exp\left(\int_0^\alpha i\frac{\overbrace{\mathbb{E}[Me^{i\nu N}\,|\,z]}^{\frac{\partial}{\partial v}\psi_{M,N|z}(v,\nu)\big|_{v=0}}}{\underbrace{\mathbb{E}[e^{i\nu N}\,|\,z]}_{\psi_{N|z}(\nu)}}d\nu\right) \qquad (1)$$

Differentiate wrt α to remove integral.

$$\frac{\frac{d}{d\alpha}\psi_{A|z}(\alpha)}{\psi_{A|z}(\alpha)} = \frac{\frac{\partial}{\partial v}\psi_{M,N|z}(v,\alpha)\Big|_{v=0}}{\psi_{N|z}(\alpha)} \qquad (2)$$

Measurement Error KIV

To obtain $\hat{\psi}^n_{A|z}$:

$$\overbrace{\mathbb{E}_{\mathscr{P}_{A|z}}[e^{i\alpha X}](\alpha)}^{\psi_{A|z}(\alpha)} = \exp\left(\int_0^\alpha i\frac{\overbrace{\mathbb{E}[Me^{i\nu N}|z]}^{\frac{\partial}{\partial v}\psi_{M,N|z}(v,\nu)\big|_{v=0}}}{\underbrace{\mathbb{E}[e^{i\nu N}|z]}_{\psi_{N|z}(\nu)}}d\nu\right) \quad (1)$$

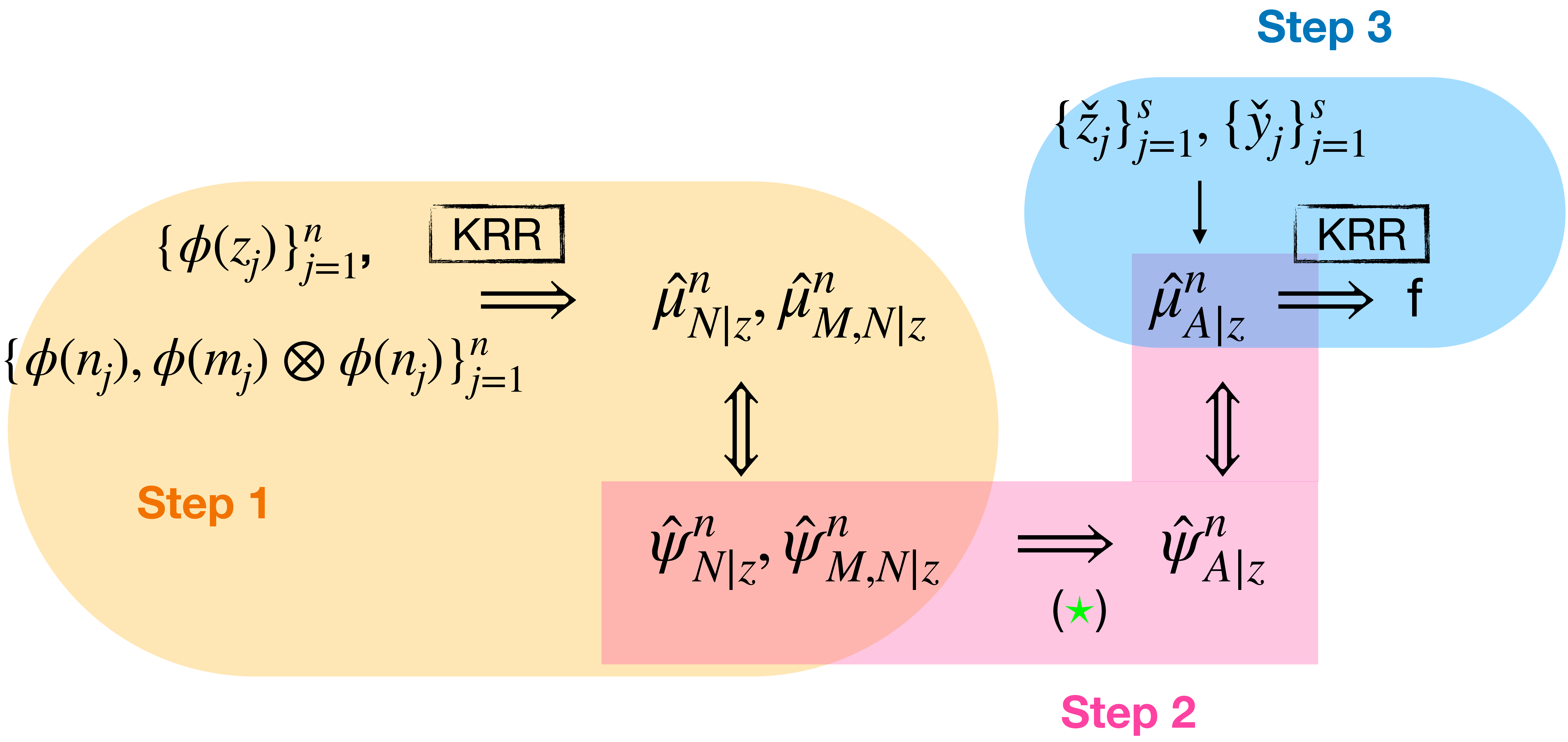
Differentiate wrt α to remove integral.

$$\frac{\frac{d}{d\alpha}\hat{\psi}^n_{A|z}(\alpha)}{\hat{\psi}^n_{A|z}(\alpha)} = \frac{\frac{\partial}{\partial v}\hat{\psi}^n_{M,N|z}(v,\alpha)\Big|_{v=0}}{\hat{\psi}^n_{N|z}(\alpha)} \quad (2)$$

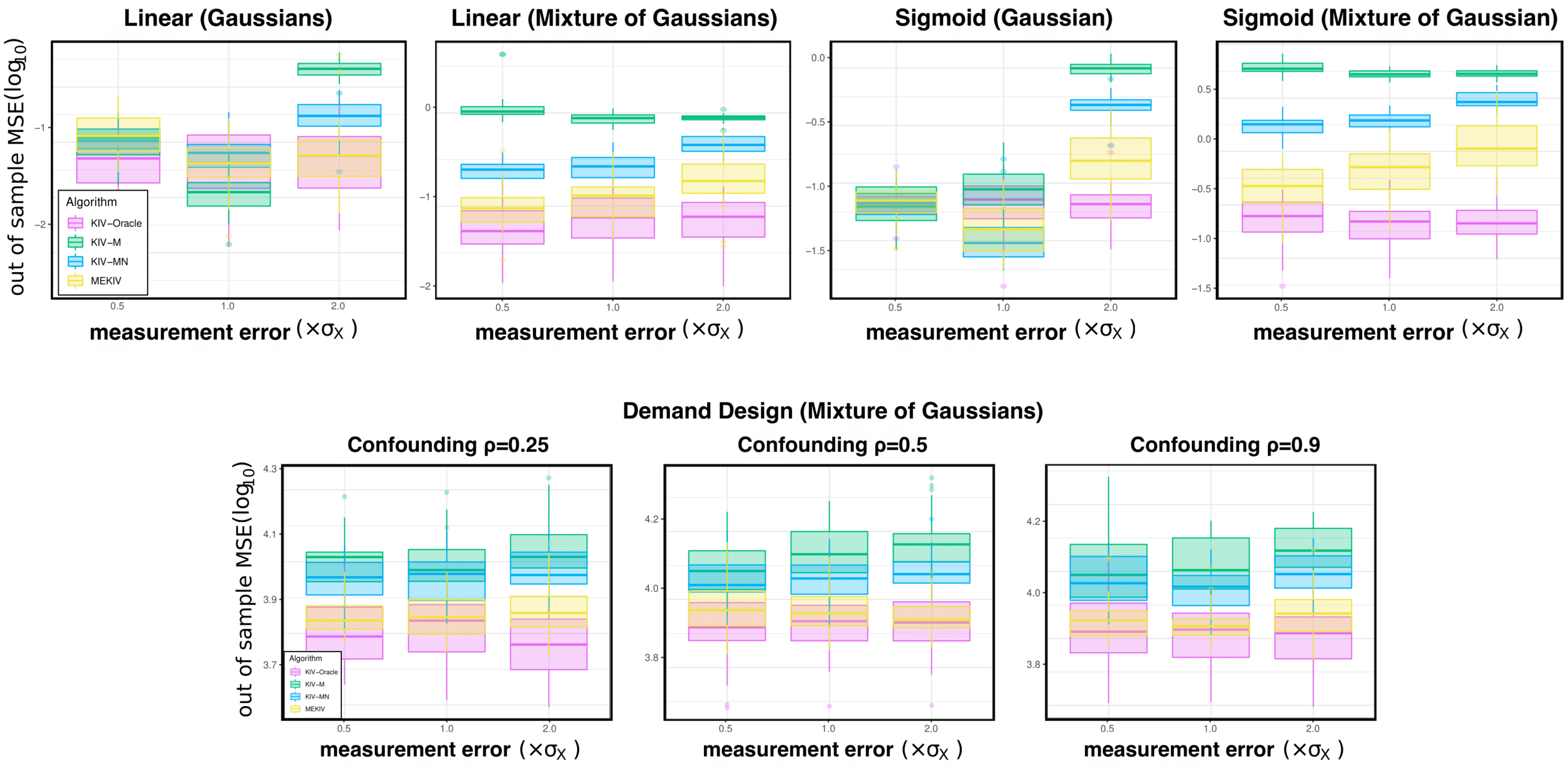
(Replace with sample estimates.)

Measurement Error KIV

Measurement Error KIV



MEKIV results



Open questions

Open questions

- Relax the measurement error assumption and IV assumption.

Open questions

- Relax the measurement error assumption and IV assumption.
- Extend to sequential settings.

Open questions

- Relax the measurement error assumption and IV assumption.
- Extend to sequential settings.

Thanks for listening! :)