Engineering Patterns in Causal Inference

Jeffrey Wong, Airbnb
Antoine Creux, Airbnb

Causal Inference is a mathematically heavy field.

$$Y = \alpha + f(x)\beta_1 + g(A)\beta_2 + (f(x) \cdot g(A))\beta_3 + \varepsilon$$

$$ATE = E[Y|A = 1] - E[Y|A = 0]$$

$$\tau(x) = E[Y|A = 1, x] - E[Y|A = 0, x]$$

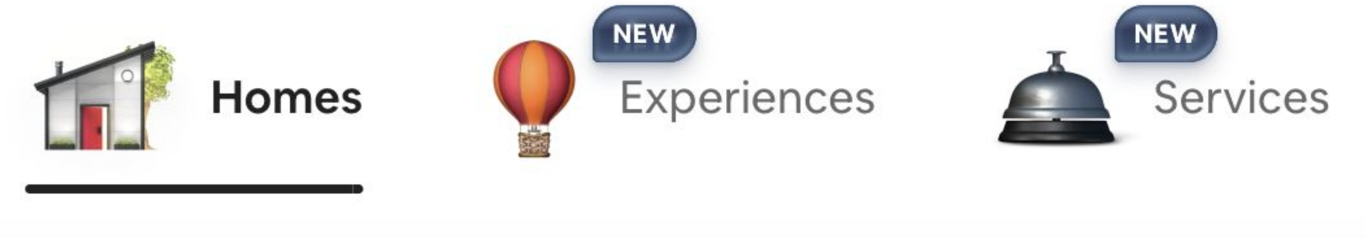
$$P(good) = Prob(\tau(x) \geq \epsilon)$$

What role does software play?

Engineering Patterns in Causal Inference

Jeffrey Wong, Airbnb
Antoine Creux, Airbnb

Causal reasoning is a principle of intelligent engineering systems.



Causal reasoning is a principle of intelligent engineering systems.

Causal reasoning is a principle of intelligent engineering systems.

	Business	Leisure
Arrived	🎈	🎈
Not Arrived	🏡	🎈

Software solves for

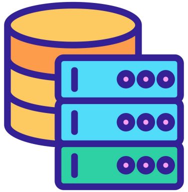
1. Scalability
2. Multiplicity
3. Grammar

Scalability

Through Numerical Methods

We need scalability at the speed of thought.

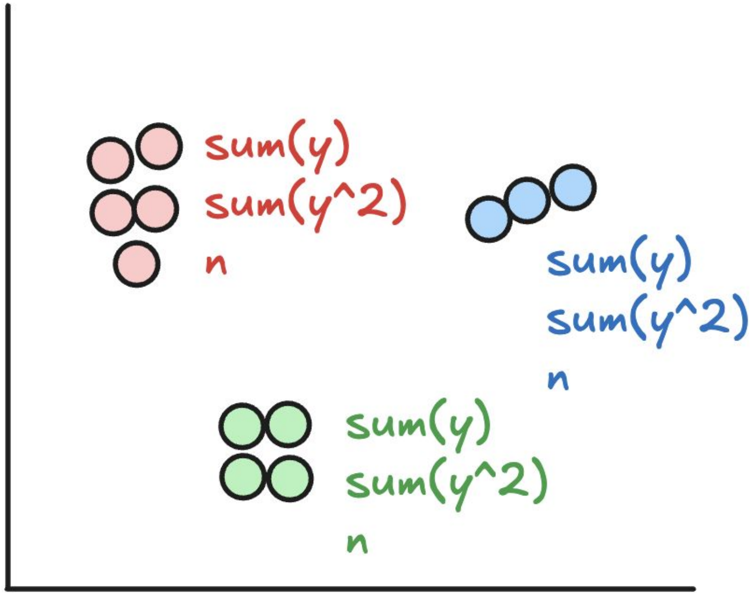
We need software that processes data and train models efficiently.



pandas

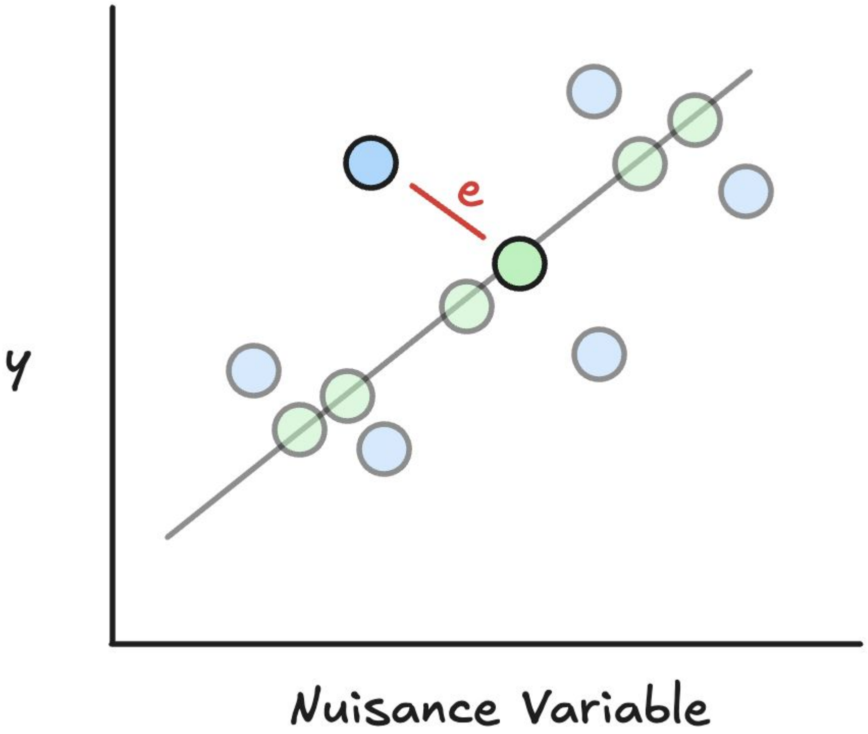
statsmodels

Sufficient Statistics lets you train on aggregates without bias.



$$
\begin{aligned}
RMSE &= \sum_i (y_i - \hat{y}_i)^2 \\
&= \sum_i (y_i^2 - 2\hat{y}_i y_i + \hat{y}_i^2) \\
&= \sum_g \sum_i (y_i^2) - \sum_g 2\hat{y}_g \sum_i y_i + \sum_g n_g \hat{y}_g^2
\end{aligned}
$$

Increase efficiency by residualizing out nuisance variables.



$$y = f(T) + g(S) + h(T, S)$$

[Quasi Oracle Estimation Of HTE (Nie and Wager)]

Querying for Multiplicity

Multiple…

1. Metrics
2. Arms
3. Segments
4. Models

Multiplicity is layers and layers of loops.

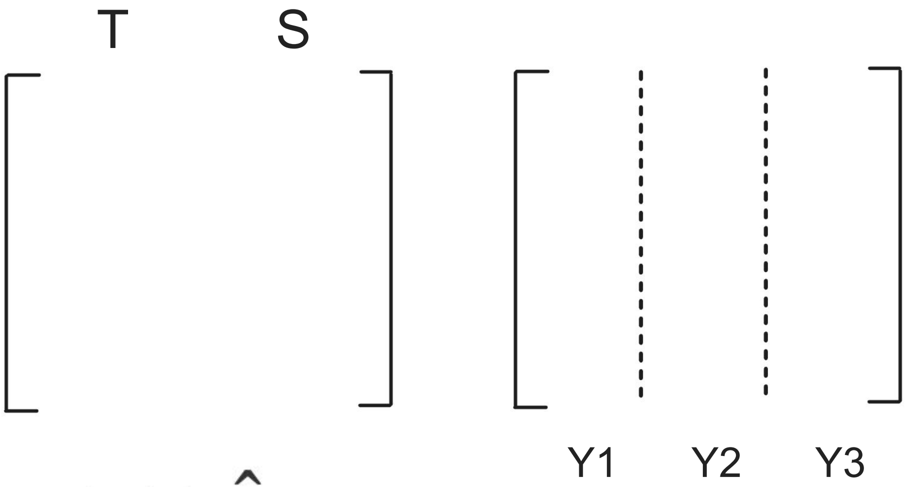
```
for model in models:
  model.fit(metrics, arms, segments)
  for metric in metrics:
    for arm in arms:
      for segment in segments:
        y_treatment = model.predict(metric, arm, segment)
        y_control   = model.predict(metric, "control", segment)
        effect = mean(y_treatment) - mean(y_control)
```

Vectorization unrolls the loops.

```
for model in models:
  model.fit(metrics, arms, segments).infer_effect(segments)
```

Vectorizing OLS

$$\hat{\beta} = (X^T X)^{-1} X^T y$$



$$\tau(s) = \Delta(s)\hat{\beta}$$

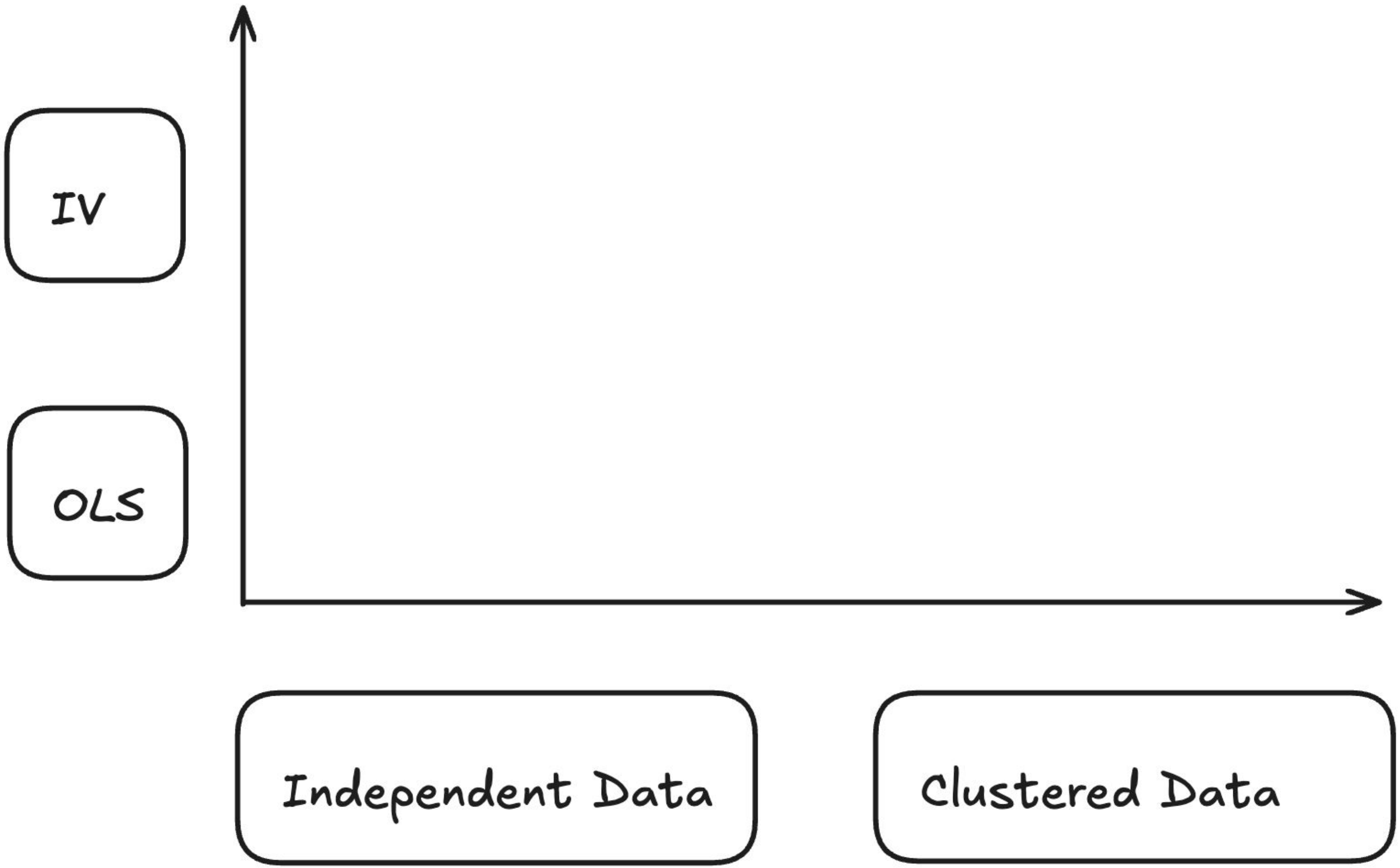
[Delta Vectors (Wong)]

Grammar

All causal inquiries have …

1. Causal identification
2. Units
3. Features & metrics
4. Model parameters
5. Inference
6. External validity
7. Policy making & evaluation

```
Data(
  treatment = "treatment",
  instrument = "treatment",
  unit_of_randomization = "host"
)
.add_metrics([
  Revenue,
  Engagement,
  unit_of_analysis = "listing"
])
```


IV
OLS
Independent Data
Clustered Data

```
.add_model(
  OLS,
  features = [...]
)
.aggregate()
```

```
.fit(
  ~ treatment + <features>
)
.infer_effect(
  condition_on = [
    "country == 'US'"
  ]
)
```

```
Data(
  treatment = "treatment",
  instrument = "treatment",
  unit_of_randomization = "host"
)
.add_metrics([
  Revenue,
  Engagement,
  unit_of_analysis = "listing"
])
.add_model(
  OLS,
  features = [...]
)
.aggregate()
.fit(
  ~ treatment + <features>
)
.infer_effect(
  condition_on = [
    "country == 'US'"
  ]
)
```

Key Takeaways

- Causal reasoning is a principle for intelligent engineering systems.
- We need engineers to think about scalable training for causal models
- Good causal software creates a stack for training models and querying models, and needs a grammar to solve for multiplicity.

If you'd like to develop together, reach out!

jeffrey.wong@airbnb.com