Tutorial:

Causal Model Search

Richard Scheines

Carnegie Mellon University

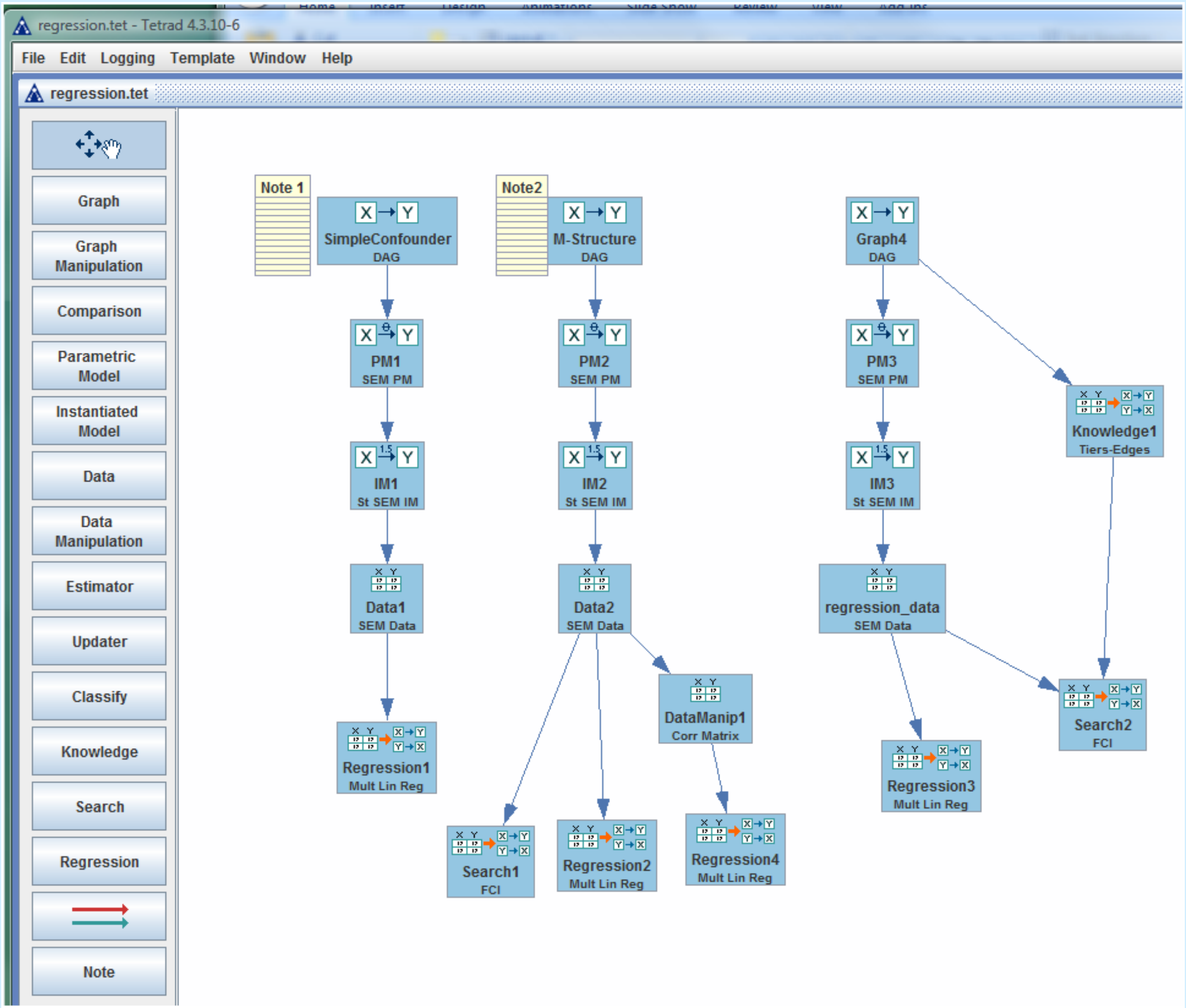
Peter Spirtes, Clark Glymour, Joe Ramsey,

others

Goals

1) Convey rudiments of graphical causal models
2) Basic working knowledge of Tetrad IV

Tetrad: Complete Causal Modeling Tool



Tetrad

1) Main website: http://www.phil.cmu.edu/projects/tetrad/
2) Download: http://www.phil.cmu.edu/projects/tetrad/current.html
3) Data files: workshop.new.files.zip in

   www.phil.cmu.edu/projects/tetrad_download/download/workshop/Data/

4) Download from Data directory:
   - tw.txt
   - Charity.txt
   - Optional:
     - estimation1.tet, estimation2.tet
     - search1.tet, search2.tet, search3.tet

Outline

1) Motivation
2) Representing/Modeling Causal Systems
3) Estimation and Model fit
4) Causal Model Search

Statistical Causal Models: Goals

1) Policy, Law, and Science: How can we use data to answer

   a) *subjunctive* questions (effects of future policy interventions), or

   b) *counterfactual* questions (what would have happened had things been done differently (law)?

   c) *scientific* questions (what mechanisms run the world)

2) Rumsfeld Problem: Do we know what we do and don't know: Can we tell when there is or is not enough information in the data to answer causal questions?

Causal Inference Requires More than Probability

Prediction from Observation ≠ Prediction from Intervention

P(**Lung Cancer 1960 = y | Tar-stained fingers 1950 = no**)

≠

P(**Lung Cancer 1960 = y | Tar-stained fingers 1950**$_{set}$ **= no**)

In general: $P(\mathbf{Y{=}y} \mid X{=}x, \mathbf{Z{=}z}) \neq P(\mathbf{Y{=}y} \mid \mathbf{X}_{set}{=}x, \mathbf{Z{=}z})$

Causal Prediction vs. Statistical Prediction:

Non-experimental data (observational study)

Background Knowledge

$P(Y,X,Z)$

Causal Structure

$P(Y{=}y \mid X{=}x, Z{=}z)$

$P(Y{=}y \mid X_{set}{=}x, Z{=}z)$

Causal Search

Causal Search:

1. Find/compute *all* the causal models that are indistinguishable given background knowledge and data
2. Represent features common to all such models

Multiple Regression is often the *wrong* tool for Causal Search:

Example: Foreign Investment & Democracy

Foreign Investment

Does Foreign Investment in 3rd World Countries *inhibit* *Democracy?*

Timberlake, M. and Williams, K. (1984). Dependence, political exclusion, and government repression: Some cross-national evidence. American Sociological Review 49, 141-146.

$N = 72$

PO	degree of political exclusivity
CV	lack of civil liberties
EN	energy consumption per capita (economic development)
FI	level of foreign investment

Foreign Investment

Correlations

	po	fi	en	cv
po	1.0			
fi	-.175	1.0		
en	-.480	0.330	1.0	
cv	0.868	-.391	-.430	1.0

Case Study: Foreign Investment

Regression Results

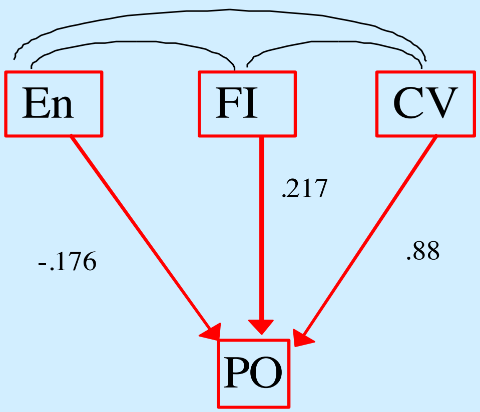
$$po = .227*fi - .176*en + .880*cv$$

	fi	en	cv
SE	(.058)	(.059)	(.060)
t	3.941	-2.99	14.6

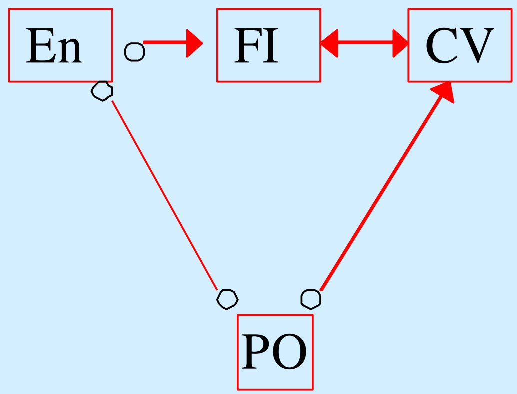
Interpretation: foreign investment increases political repression

Case Study: Foreign Investment

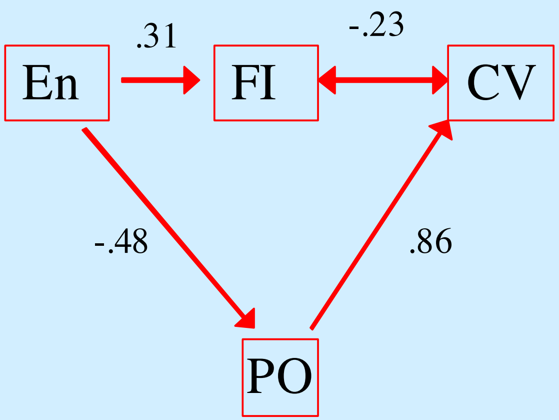
Alternatives



Regression



Tetrad - FCI



Fit: df=2, $\chi 2$=0.12, p-value = .94

There is no model with testable constraints (df > 0) that is not rejected by the data, in which FI has a positive effect on PO.

Tetrad Demo

1. Load tw.txt data
2. Estimate regression
3. Search for alternatives
4. Estimate alternative

Tetrad Hands-On

1. Load tw.txt data
2. Estimate regression

Outline

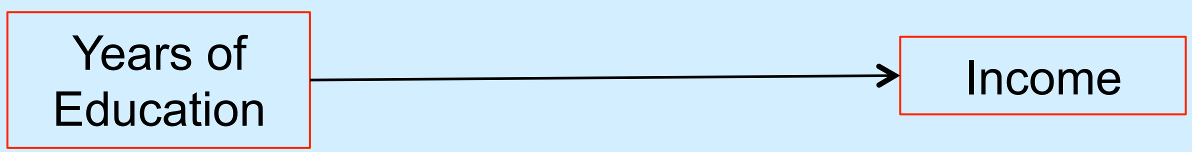
1) Motivation
2) Representing/Modeling Causal Systems
    1) Causal Graphs
    2) Standard Parametric Models
        1) Bayes Nets
        2) Structural Equation Models
    3) Other Parametric Models
        1) Generalized SEMs
        2) Time Lag models

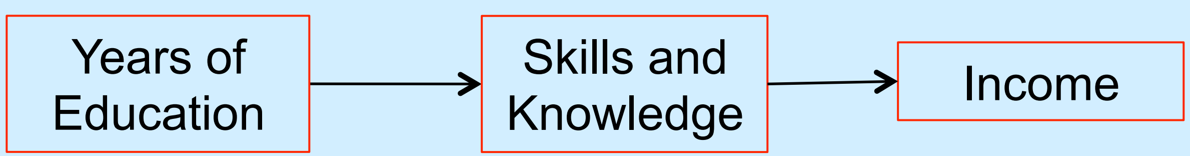
Causal Graphs

Causal Graph G = {**V**,**E**}

Each edge X → Y represents a direct causal claim:

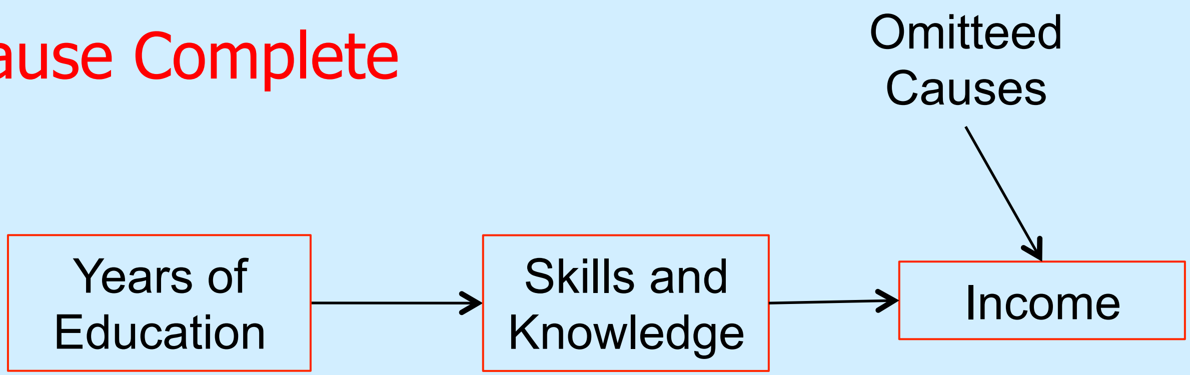
X is a direct cause of Y relative to **V**



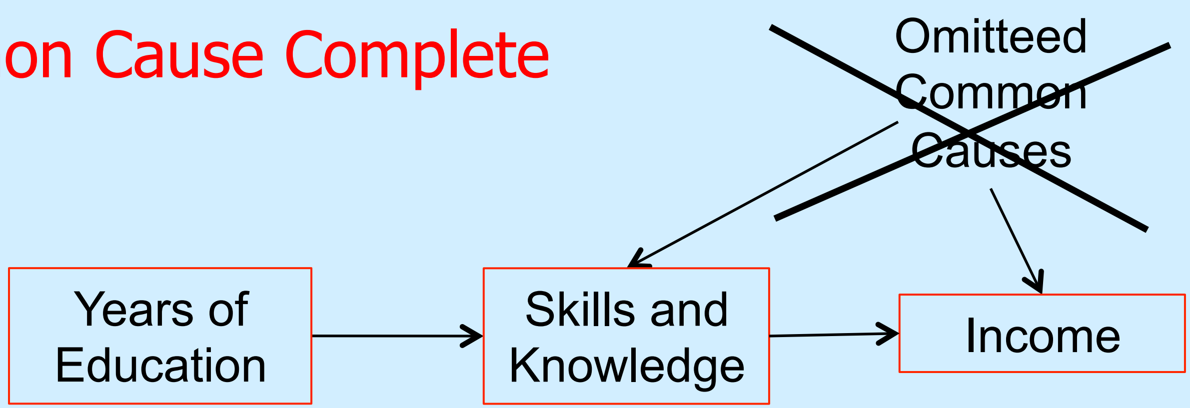


Causal Graphs

Not Cause Complete



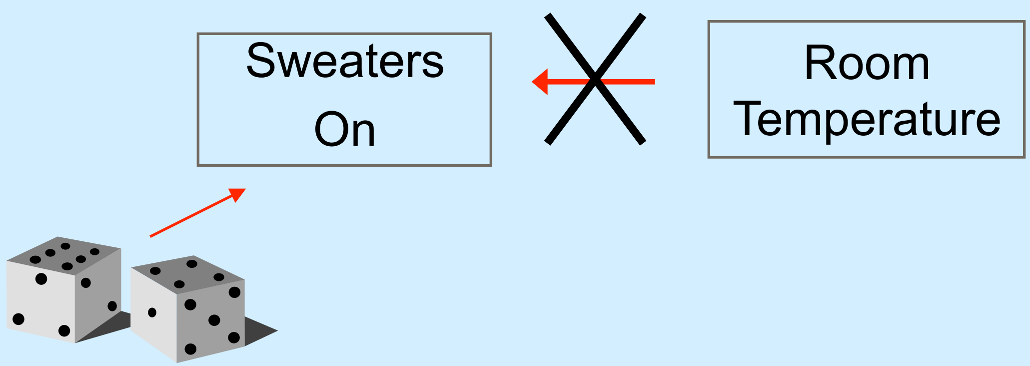
Common Cause Complete



Modeling Ideal Interventions

Interventions on the Effect

Post experimental System



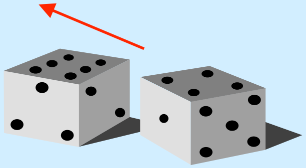
Modeling Ideal Interventions

Interventions on the Cause

Post-experimental System

Sweaters On

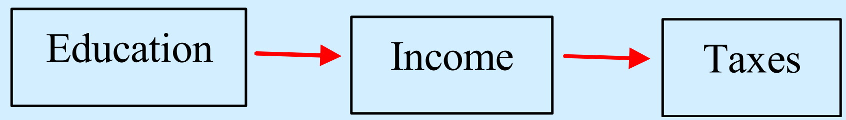
Room Temperature



Interventions & Causal Graphs

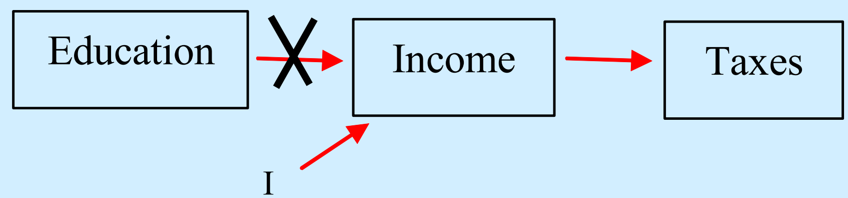
Model an ideal intervention by adding an "intervention" variable outside the original system as a direct cause of its target.

Pre-intervention graph

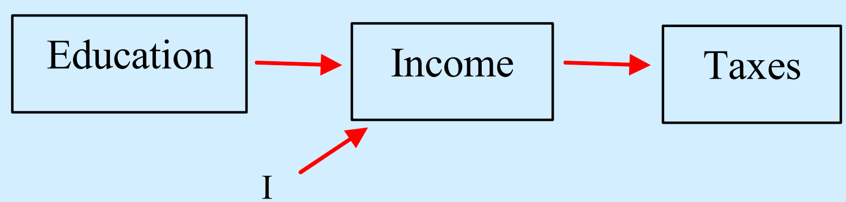


Intervene on *Income*

"Hard" Intervention



"Soft" Intervention



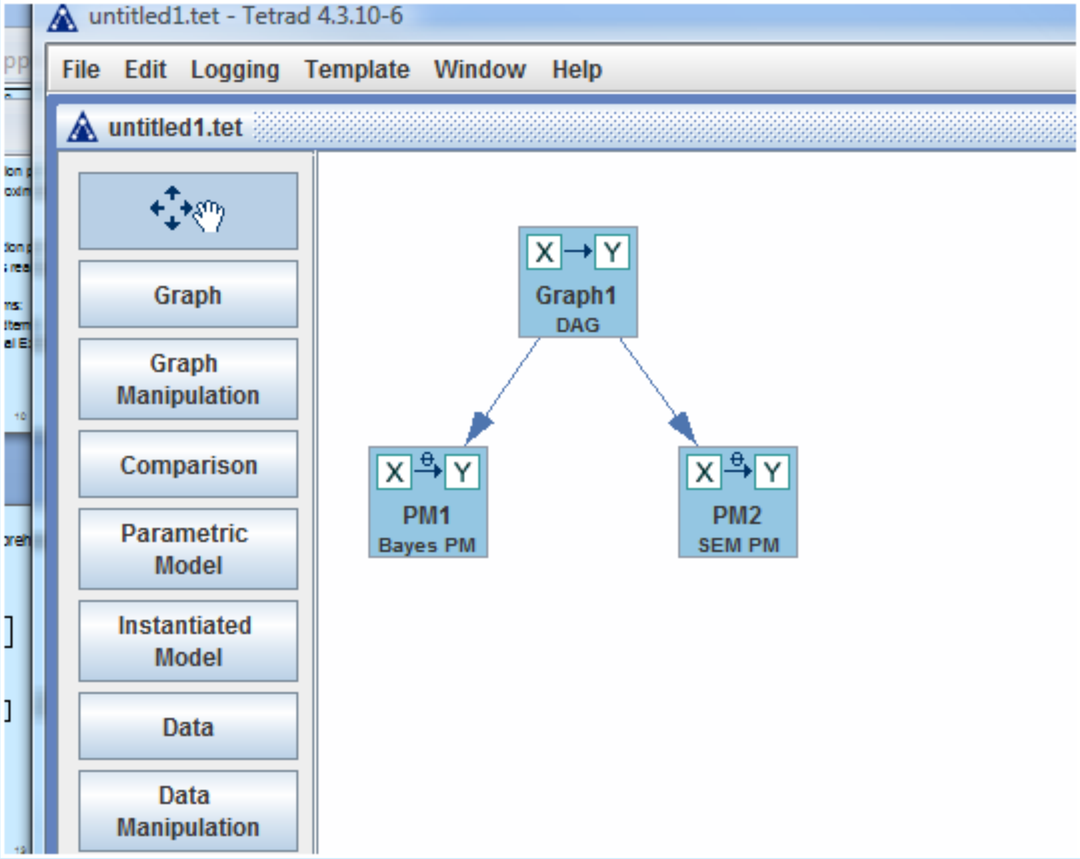
Tetrad Demo & Hands-On

Build and Save an acyclic causal graph:

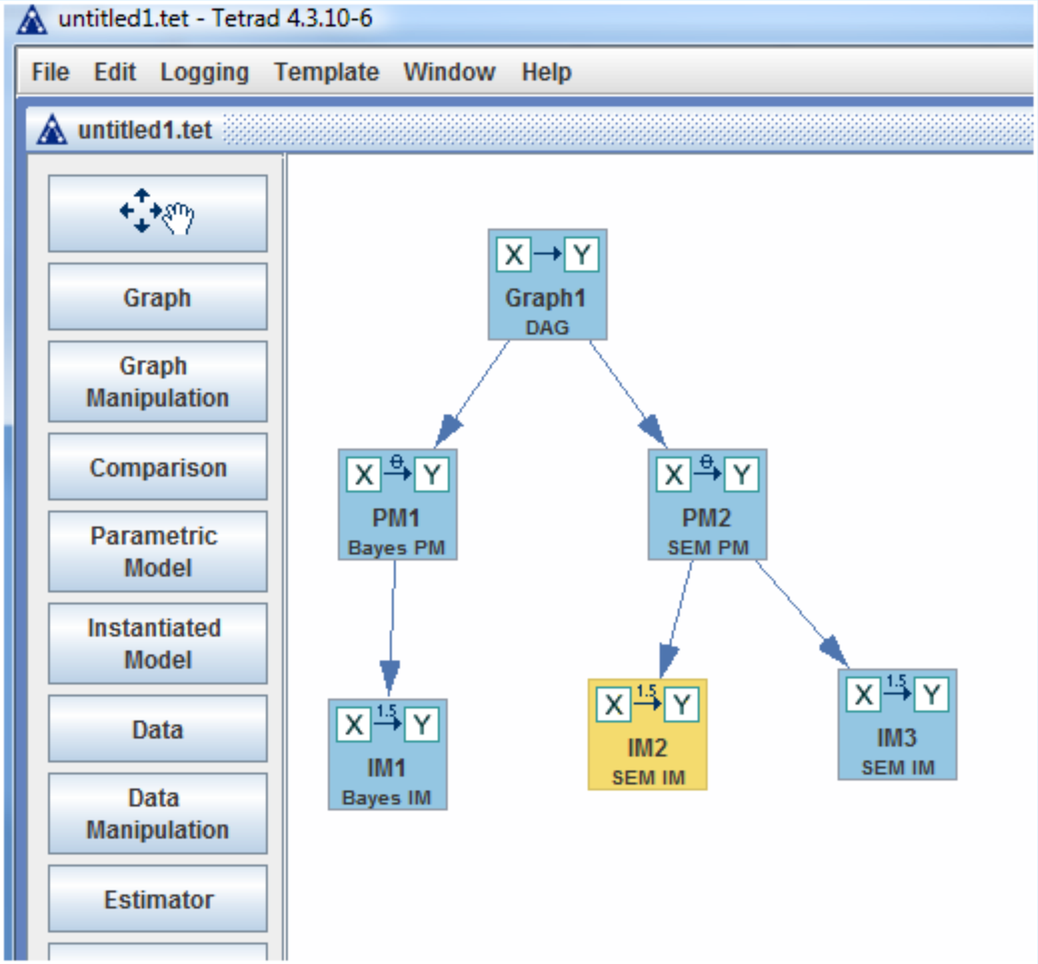
1) with 3 measured variables, no latents

2) with 5 variables, and at least 1 latent

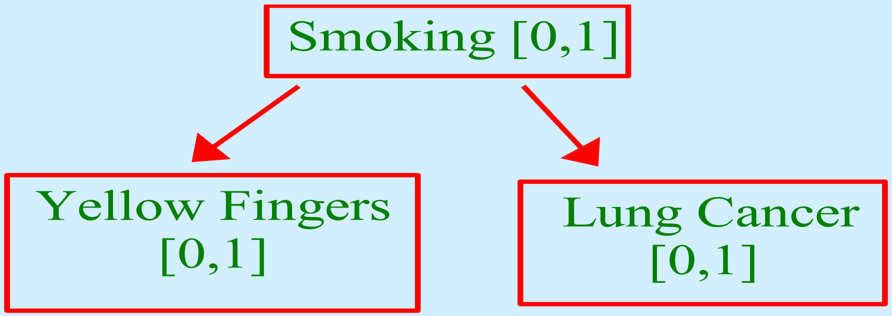
Parametric Models



Instantiated Models



Causal Bayes Networks



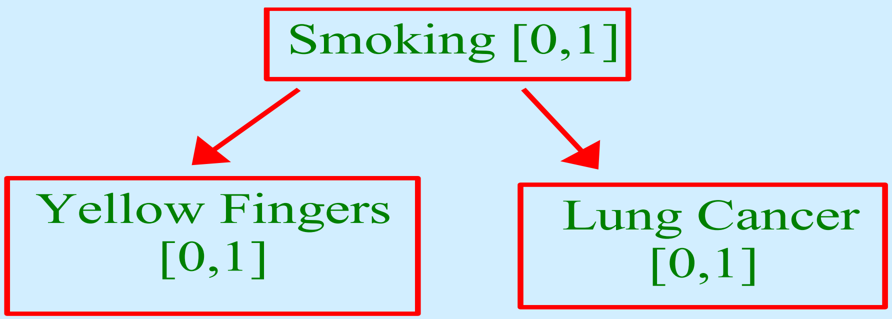
The Joint Distribution Factors

According to the Causal Graph,

$$P(V) = \prod_{x \in V} \mathbf{P}(X \mid Direct\_causes(X))$$

P(S,YF, L) = P(S) P(YF | S) P(LC | S)

Causal Bayes Networks



The Joint Distribution Factors

According to the Causal Graph,

$$P(V) = \prod_{x \in V} \mathbf{P}(X \mid Direct\_causes(X))$$

P(S) P(YF | S) P(LC | S) = f(θ)

All variables binary [0,1]: $\theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$

$P(S = 0) = \theta_1$

$P(S = 1) = 1 - \theta_1$

$P(YF = 0 \mid S = 0) = \theta_2$

$P(YF = 1 \mid S = 0) = 1 - \theta_2$

$P(YF = 0 \mid S = 1) = \theta_3$

$P(YF = 1 \mid S = 1) = 1 - \theta_3$

$P(LC = 0 \mid S = 0) = \theta_4$

$P(LC = 1 \mid S = 0) = 1 - \theta_4$

$P(LC = 0 \mid S = 1) = \theta_5$

$P(LC = 1 \mid S = 1) = 1 - \theta_5$

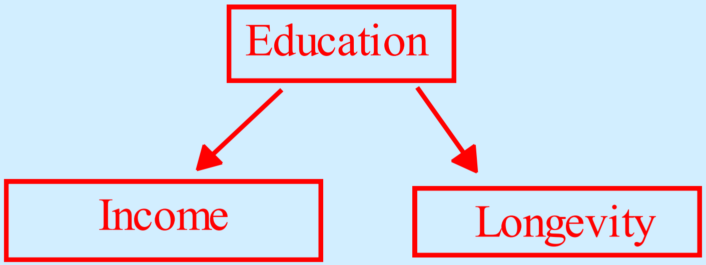
Tetrad Demo & Hands-On

1) Attach a Bayes PM to your 3-variable graph
2) Define the Bayes PM (# and values of categories for each variable)

3) Attach an IM to the Bayes PM
4) Fill in the Conditional Probability Tables.

Structural Equation Models

Causal Graph



- Structural Equations

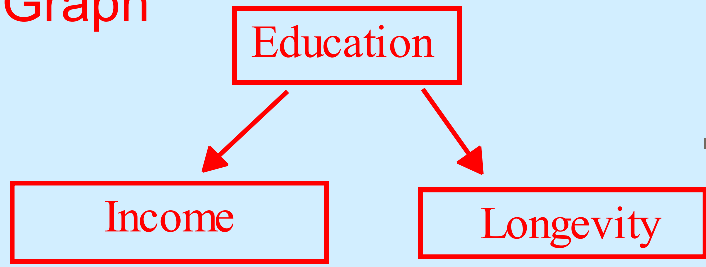
  For each variable $X \in \mathbf{V}$, an *assignment* equation:

  $$X := f_X(\text{immediate-causes}(X), \varepsilon_X)$$

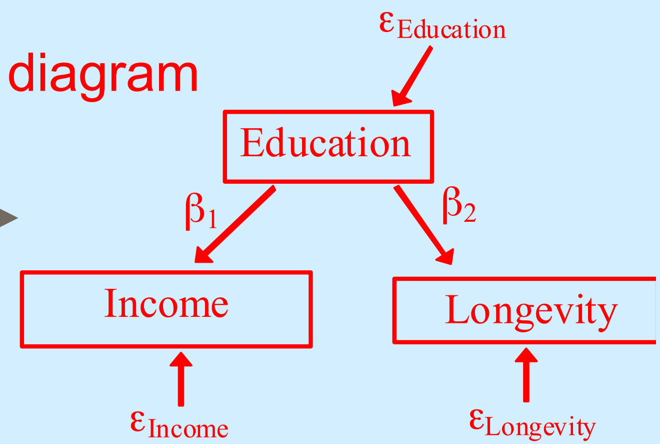
- Exogenous Distribution: Joint distribution over the exogenous vars : $P(\varepsilon)$

Linear Structural Equation Models

Causal Graph



Path diagram



Equations:

Education := $\varepsilon_{Education}$

Income := β_1 Education + ε_{income}

Longevity := β_2 Education + $\varepsilon_{Longevity}$

Structural Equation Model:

V = B**V** + **E**

Exogenous Distribution:

$P(\varepsilon_{ed}, \varepsilon_{Income}, \varepsilon_{Income})$

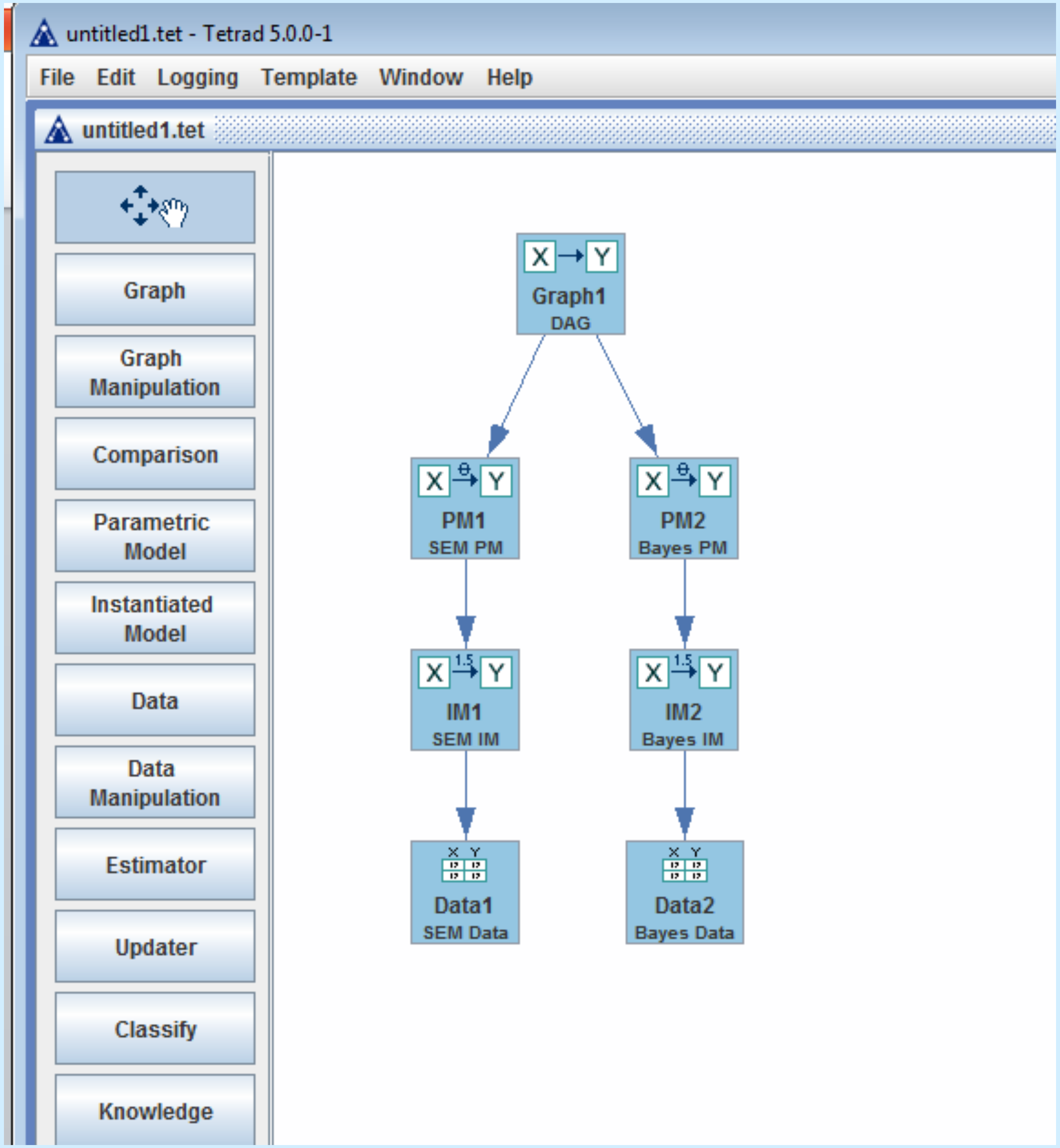
- $\forall i \neq j\ \varepsilon_i \perp \varepsilon_j$ (pairwise independence)
- no variance is zero

E.g.

$(\varepsilon_{ed}, \varepsilon_{Income}, \varepsilon_{Income}) \sim N(0, \Sigma^2)$

– Σ^2 diagonal,
- no variance is zero

Simulated Data



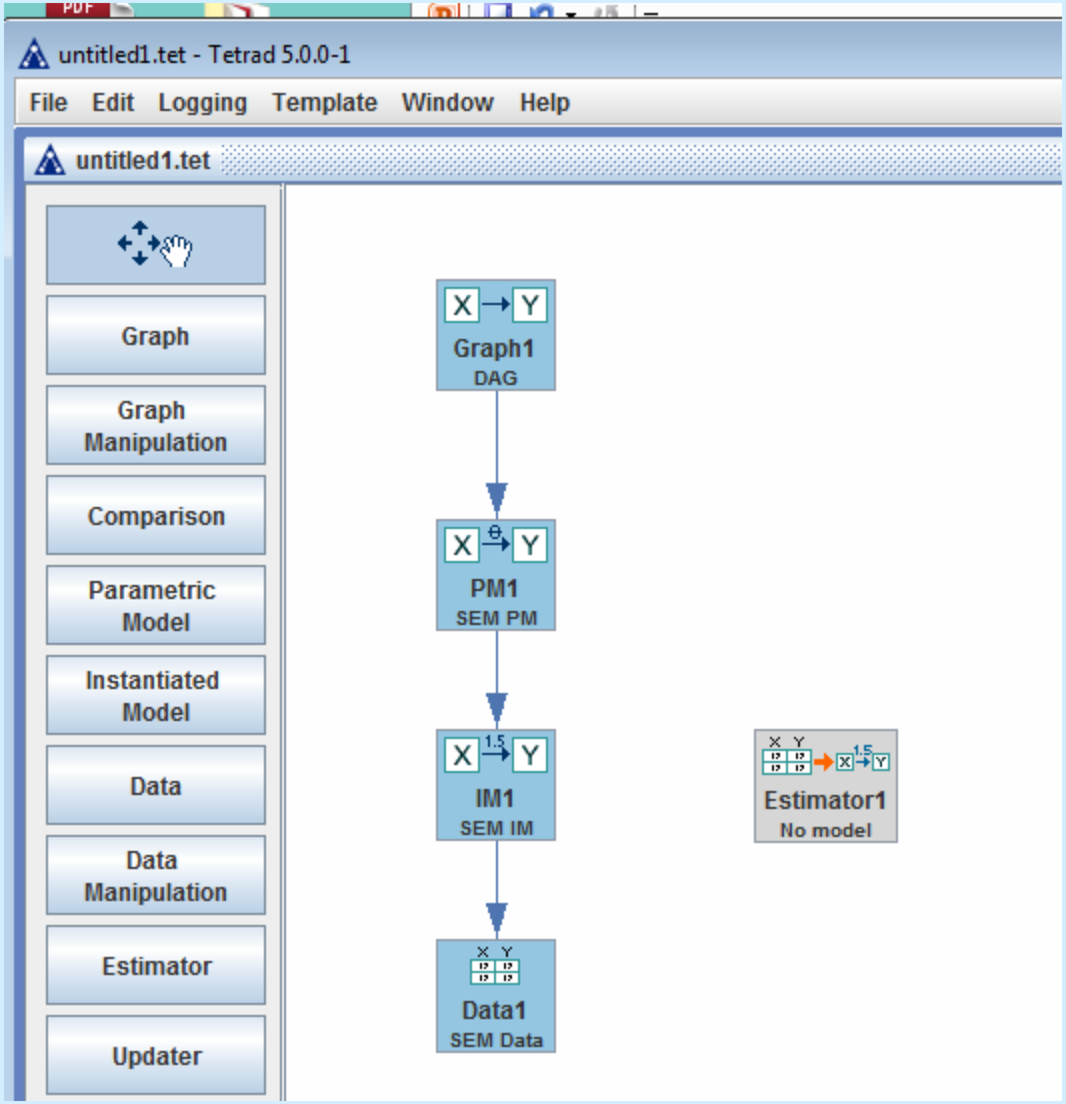
Tetrad Demo & Hands-On

1) Attach a SEM PM to your 3-variable graph
2) Attach a SEM IM to the SEM PM
3) Change the coefficient values.
4) Simulate Data from both your SEM IM and your Bayes IM

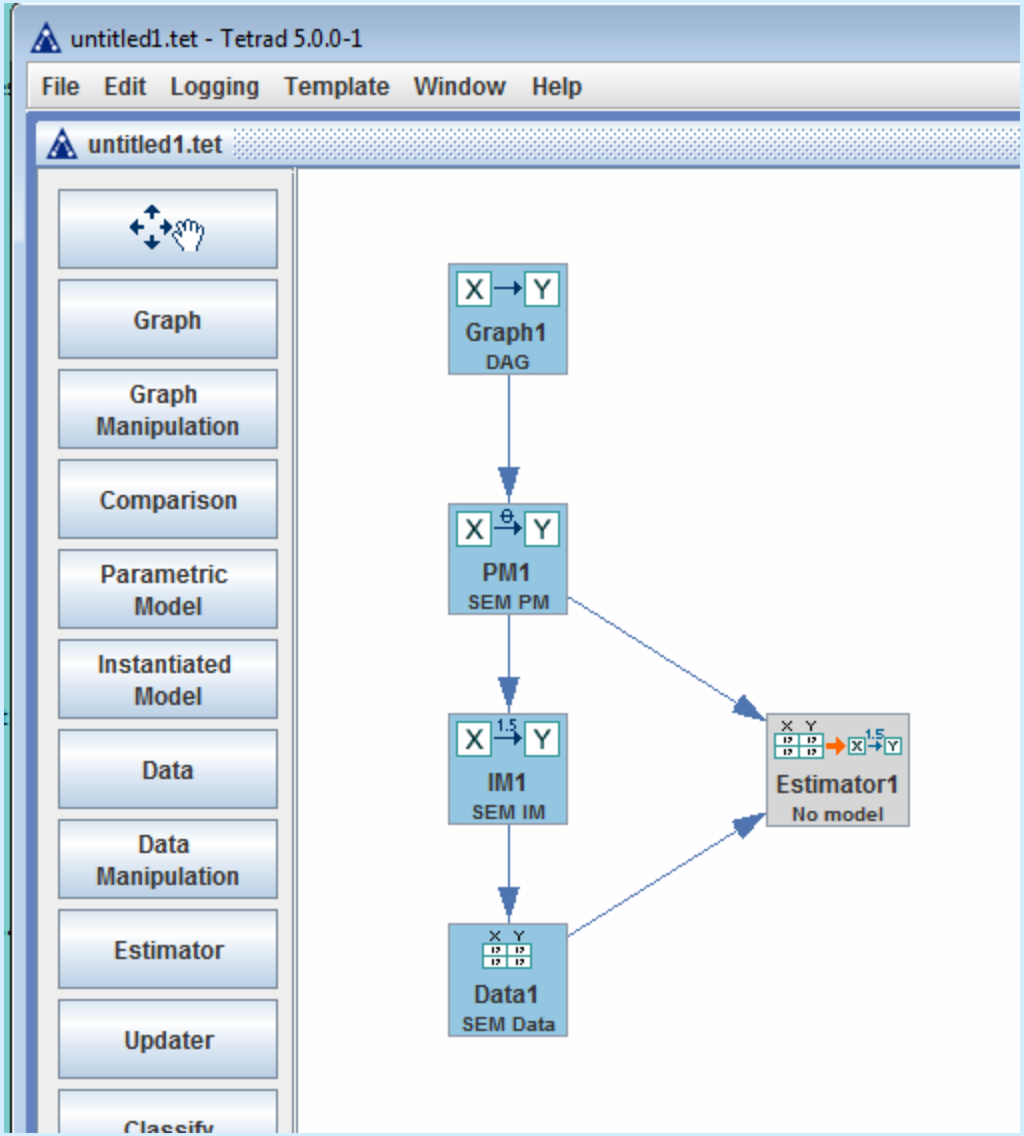
Outline

1) Motivation
2) Representing/Modeling Causal Systems
3) Estimation and Model fit
4) Model Search

Estimation


untitled1.tet - Tetrad 5.0.0-1
File Edit Logging Template Window Help
untitled1.tet
Graph
Graph Manipulation
Comparison
Parametric Model
Instantiated Model
Data
Data Manipulation
Estimator
Updater
X Y
Graph1
DAG
X θ Y
PM1
SEM PM
X 1.5 Y
IM1
SEM IM
X Y
Data1
SEM Data
X Y X 1.5 Y
Estimator1
No model

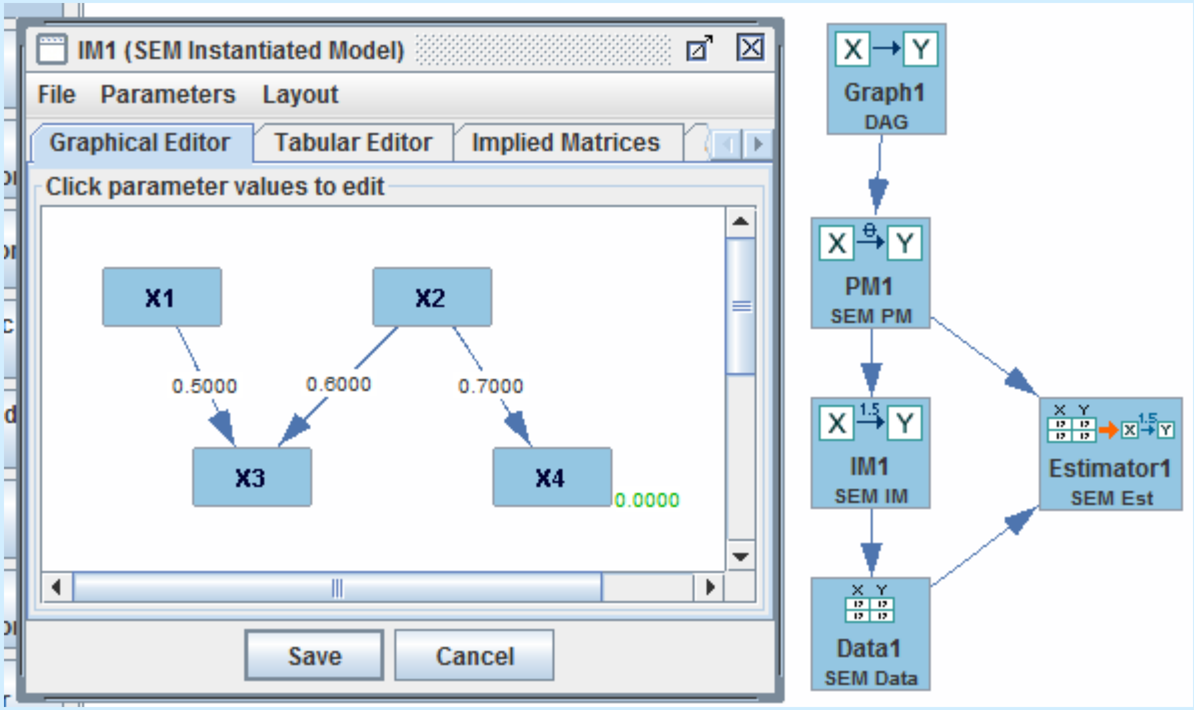
Estimation



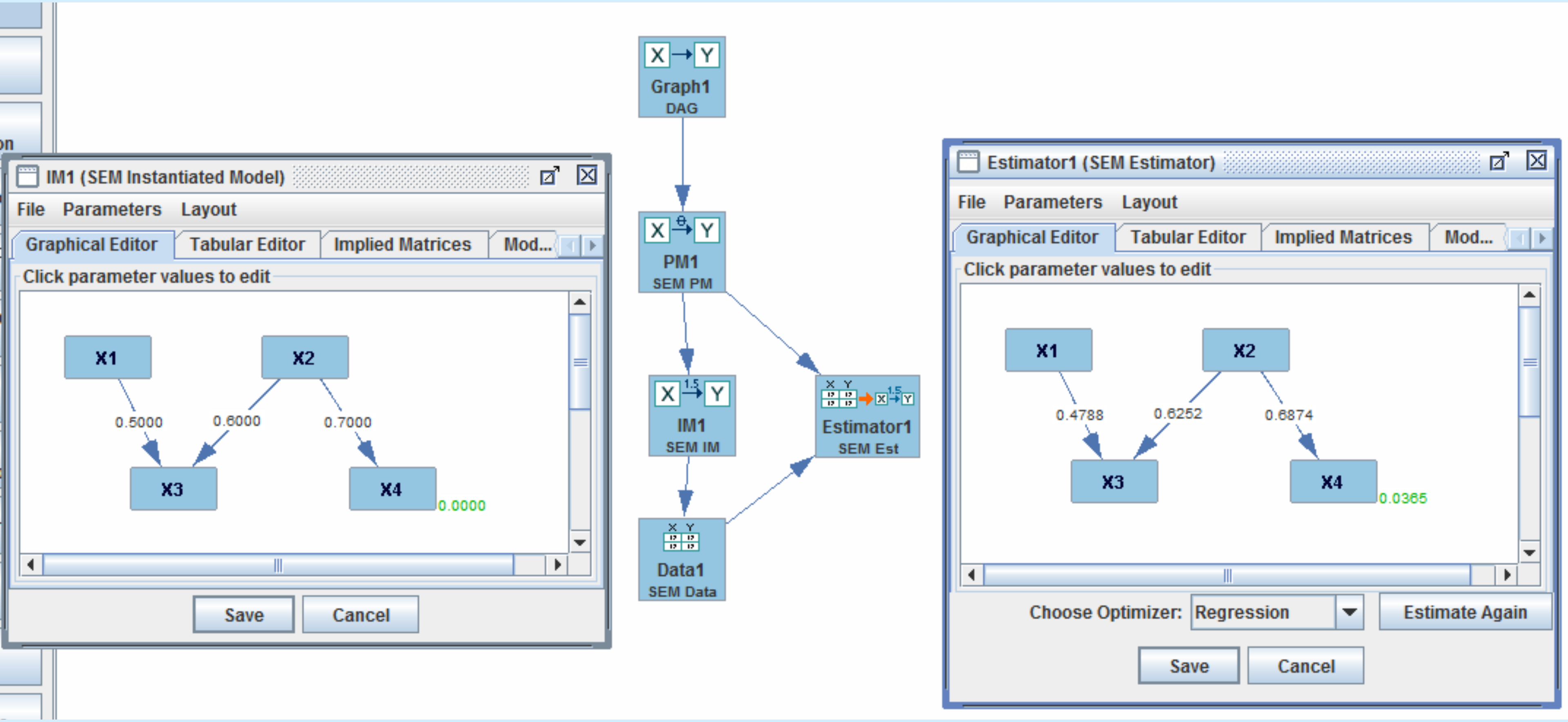
untitled1.tet - Tetrad 5.0.0-1
File Edit Logging Template Window Help
untitled1.tet
Graph
Graph Manipulation
Comparison
Parametric Model
Instantiated Model
Data
Data Manipulation
Estimator
Updater
Graph1
DAG
PM1
SEM PM
IM1
SEM IM
Data1
SEM Data
Estimator1
No model

Tetrad Demo and Hands-on

1) Select Template: “Estimate from Simulated Data”
2) Build the SEM shown below – all error standard deviations = 1.0 (go into the Tabular Editor)
3) Generate simulated data N=1000
4) Estimate model.
5) Save session as “Estimate1”



Estimation



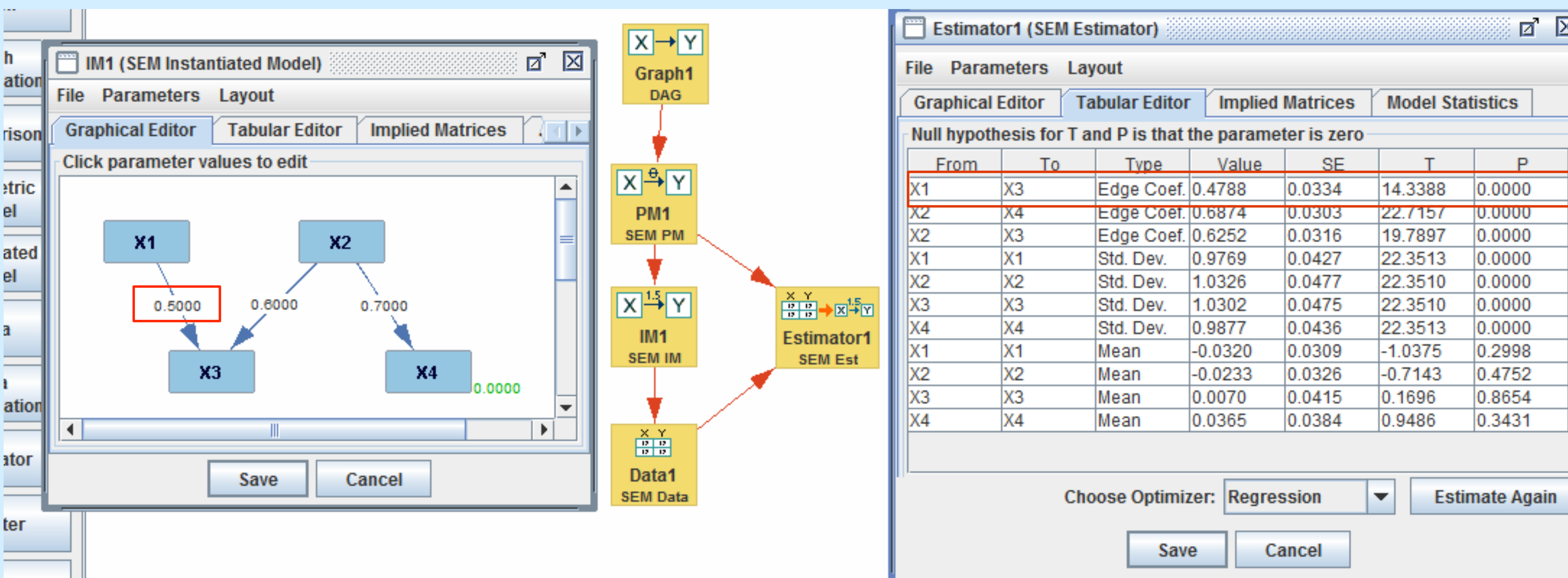
IM1 (SEM Instantiated Model)
File Parameters Layout
Graphical Editor
Tabular Editor
Implied Matrices
Mod...
Click parameter values to edit
X1
X2
X3
X4
0.5000
0.6000
0.7000
0.0000
Save
Cancel
Graph1
DAG
PM1
SEM PM
IM1
SEM IM
Data1
SEM Data
Estimator1
SEM Est
Estimator1 (SEM Estimator)
File Parameters Layout
Graphical Editor
Tabular Editor
Implied Matrices
Mod...
Click parameter values to edit
X1
X2
X3
X4
0.4788
0.6252
0.6874
0.0365
Choose Optimizer:
Regression
Estimate Again
Save
Cancel

Coefficient inference vs. Model Fit

Coefficient Inference: Null: coefficient = 0

p-value = p(Estimated value $\hat{\beta}_{X1\rightarrow X3} \geq .4788 \mid \beta_{X1\rightarrow X3} = 0$ & *rest of model correct*)

Reject null (coefficient is "significant") when p-value < α, α usually = .05



From	To	Type	Value	SE	T	P
X1	X3	Edge Coef.	0.4788	0.0334	14.3388	0.0000
X2	X4	Edge Coef.	0.6874	0.0303	22.7157	0.0000
X2	X3	Edge Coef.	0.6252	0.0316	19.7897	0.0000
X1	X1	Std. Dev.	0.9769	0.0427	22.3513	0.0000
X2	X2	Std. Dev.	1.0326	0.0477	22.3510	0.0000
X3	X3	Std. Dev.	1.0302	0.0475	22.3510	0.0000
X4	X4	Std. Dev.	0.9877	0.0436	22.3513	0.0000
X1	X1	Mean	-0.0320	0.0309	-1.0375	0.2998
X2	X2	Mean	-0.0233	0.0326	-0.7143	0.4752
X3	X3	Mean	0.0070	0.0415	0.1696	0.8654
X4	X4	Mean	0.0365	0.0384	0.9486	0.3431

Coefficient inference vs. Model Fit

Coefficient Inference: Null: coefficient = 0

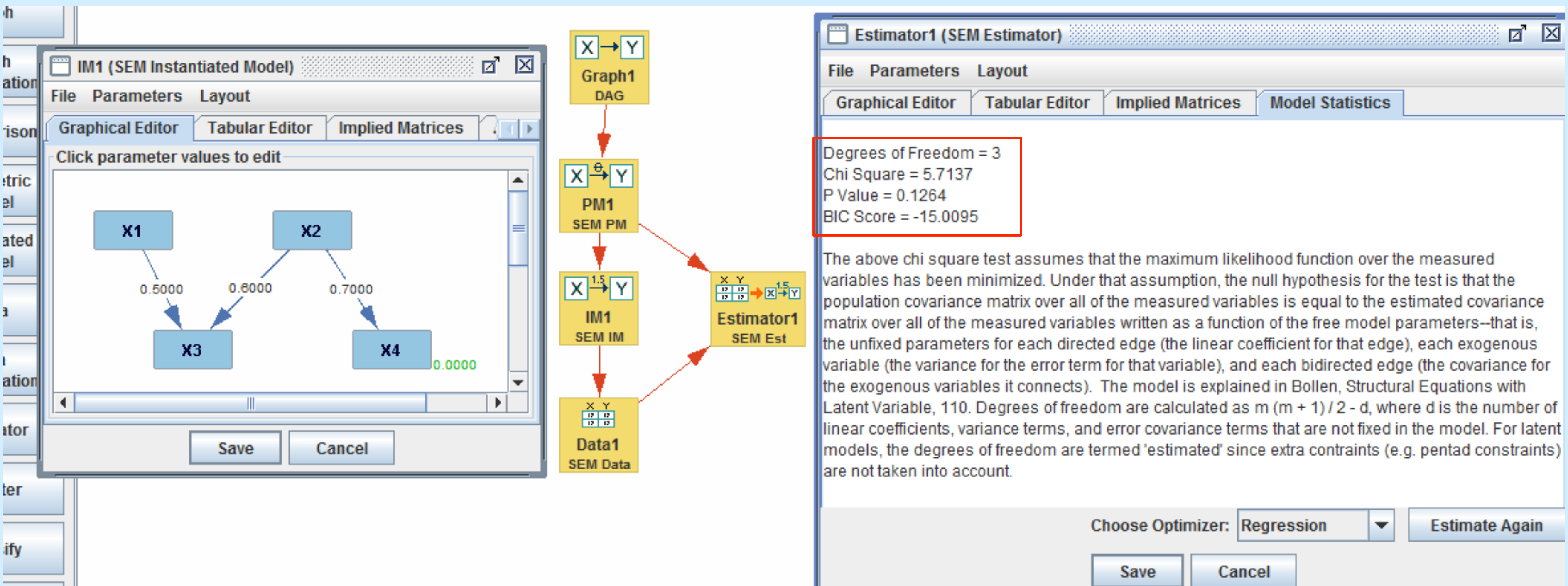
p-value = p(Estimated value $\widehat{\beta}_{X1 \rightarrow X3} \geq .4788 \mid \beta_{X1 \rightarrow X3} = 0$ & *rest of model correct*)

Reject null (coefficient is "significant") when p-value < < α, α usually = .05,

---

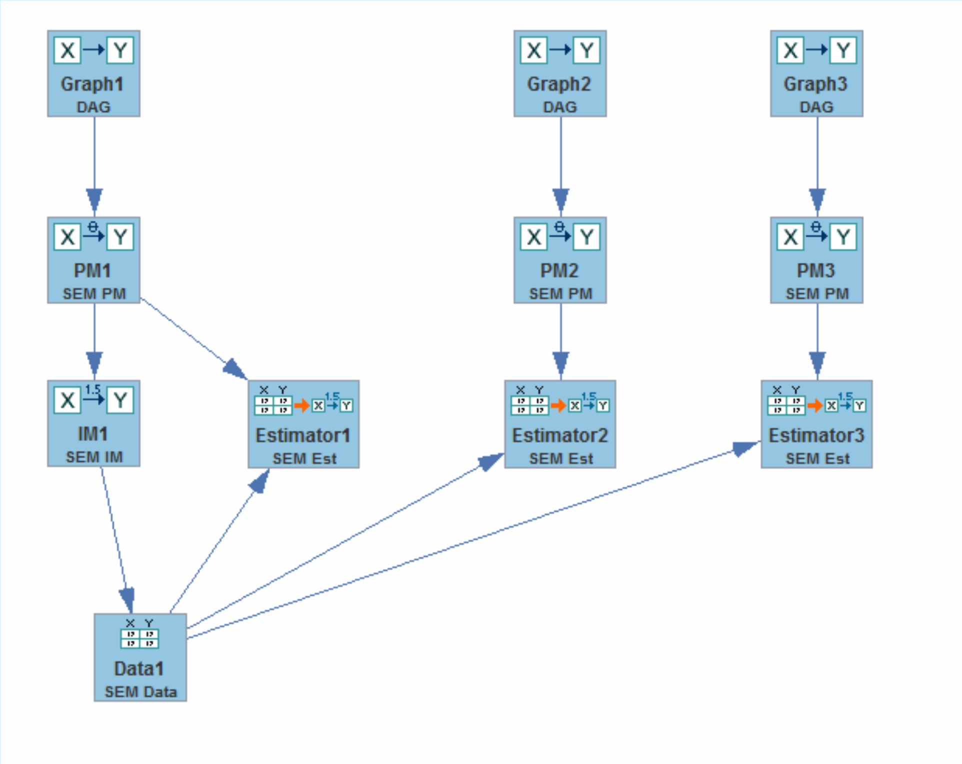
Model fit: Null: Model is correctly specified (constraints true in population)

p-value = p(f(Deviation(Σ_{ml},S)) $\geq$ 5.7137 | Model correctly specified)



Tetrad Demo and Hands-on

1) Create two DAGs with the same variables – each with one edge flipped, and attach a SEM PM to each new graph (copy and paste by selecting nodes, Ctl-C to copy, and then Ctl-V to paste)
2) Estimate each new model on the data produced by original graph
3) Check p-values of:
   a) Edge coefficients
   b) Model fit
4) Save session as: "session2"



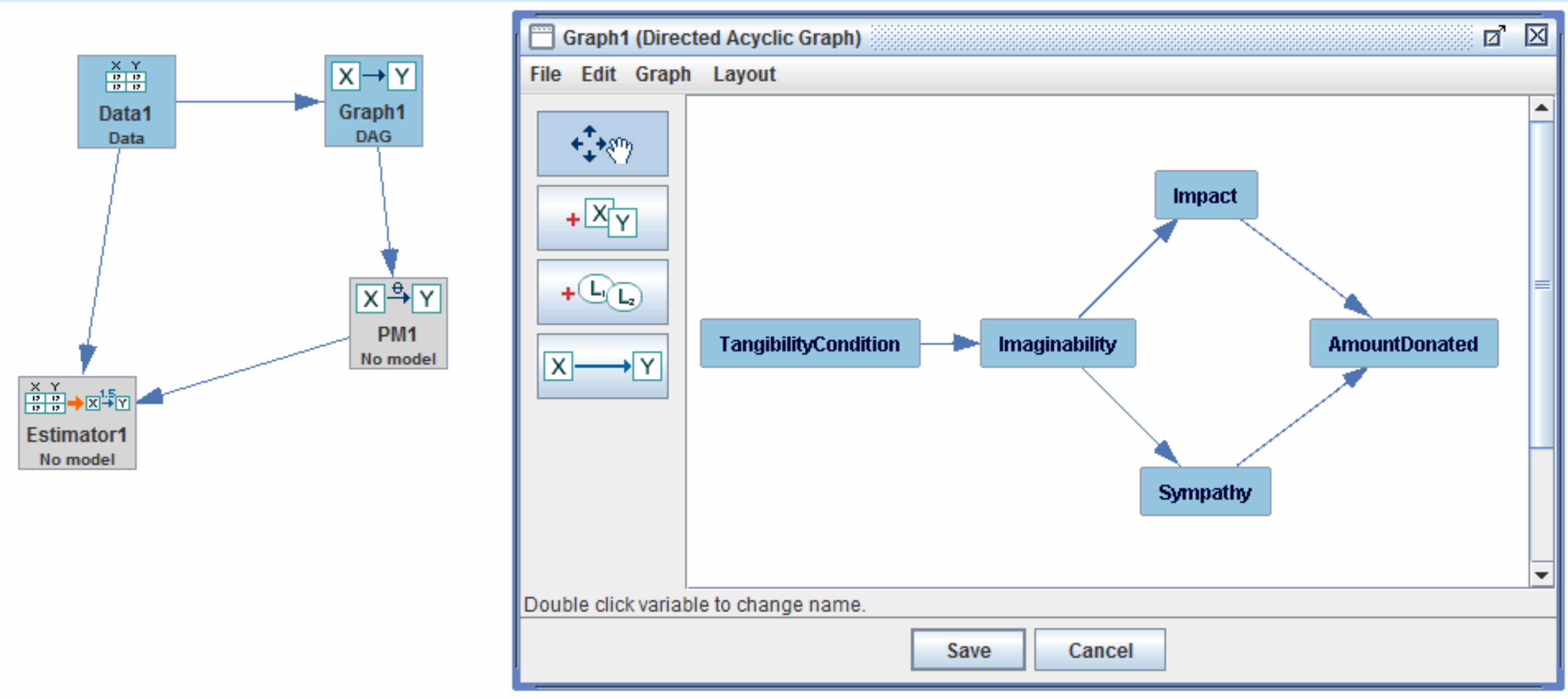
Charitable Giving

What influences giving? Sympathy? Impact?

"The Donor is in the Details", Organizational Behavior and Human Decision Processes, Issue 1, 15-23, with G. Loewenstein, R. Scheines.

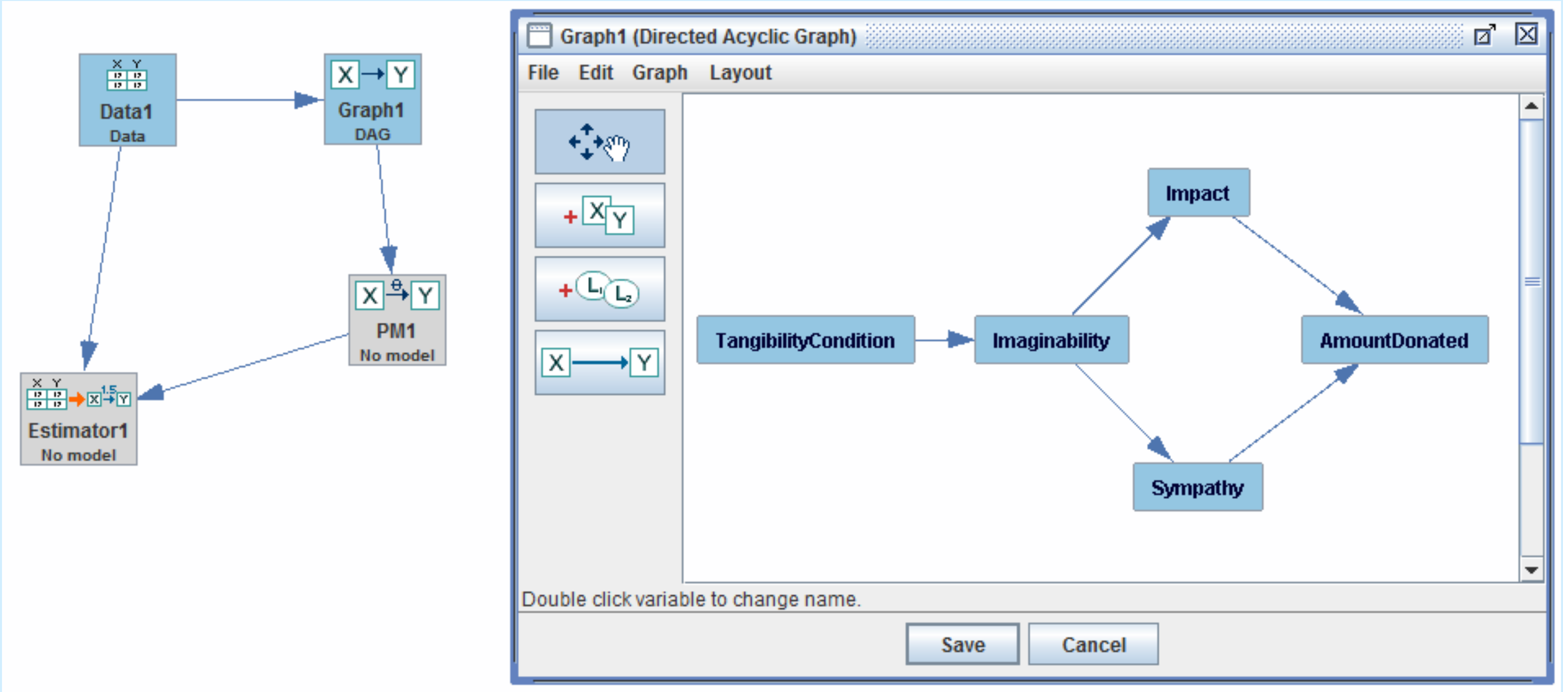
		N = 94
TangibilityCondition	[1,0]	Randomly assigned experimental condition
Imaginability	[1..7]	How concrete scenario I
Sympathy	[1..7]	How much sympathy for target
Impact	[1..7]	How much impact will my donation have
AmountDonated	[0..5]	How much actually donated

Theoretical Hypothesis



Tetrad Demo and Hands-on

1) Load charity.txt (tabular – not covariance data)
2) Build graph of theoretical hypothesis
3) Build SEM PM from graph
4) Estimate PM, check results

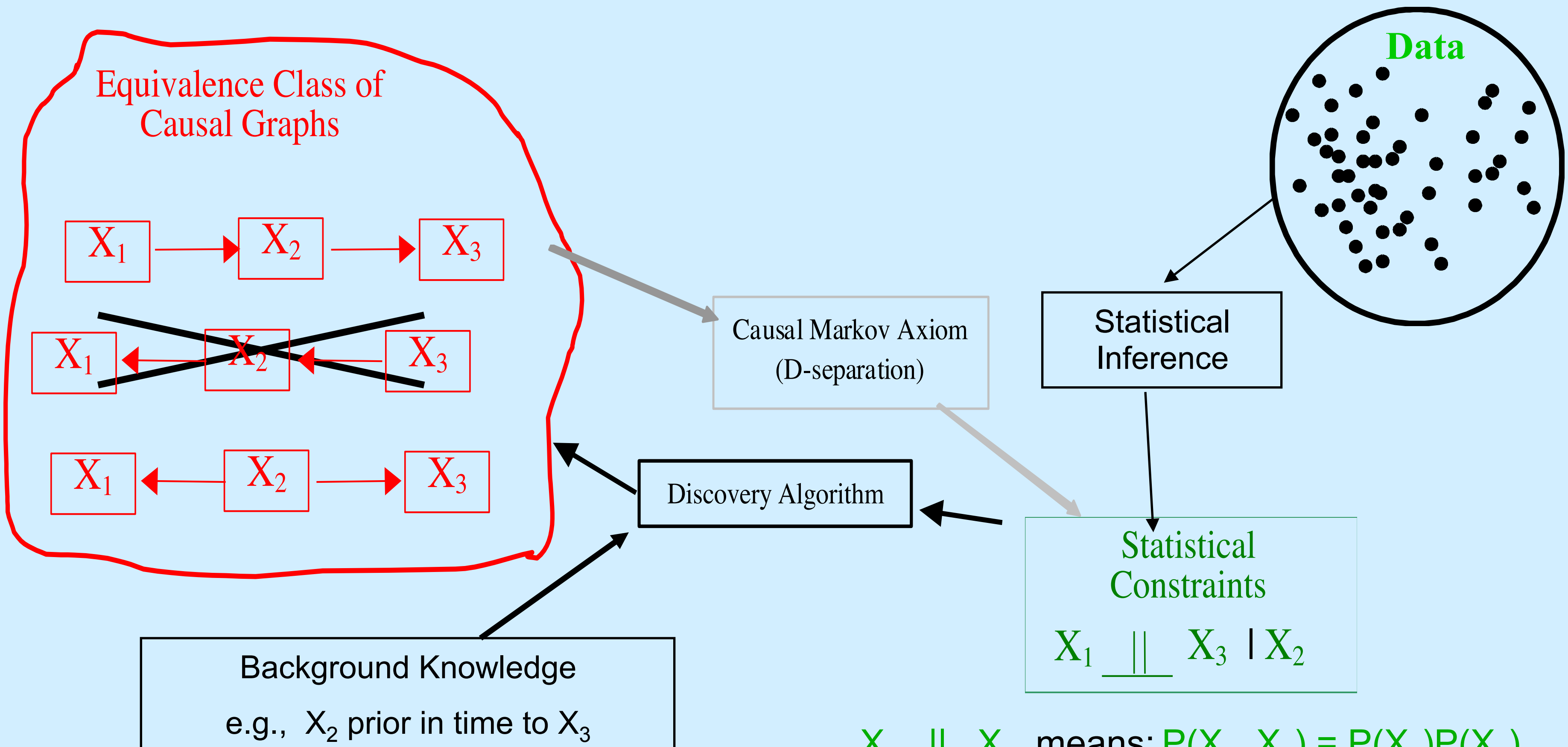


10 Minute Break

Outline

1) Motivation
2) Representing/Modeling Causal Systems
3) Estimation and Model fit
4) Model Search
    1) Bridge Principles (Causal Graphs ⇔ Probability Constraints):
        a) Markov assumption
        b) Faithfulness assumption
        c) D-separation
    2) Equivalence classes
    3) Search

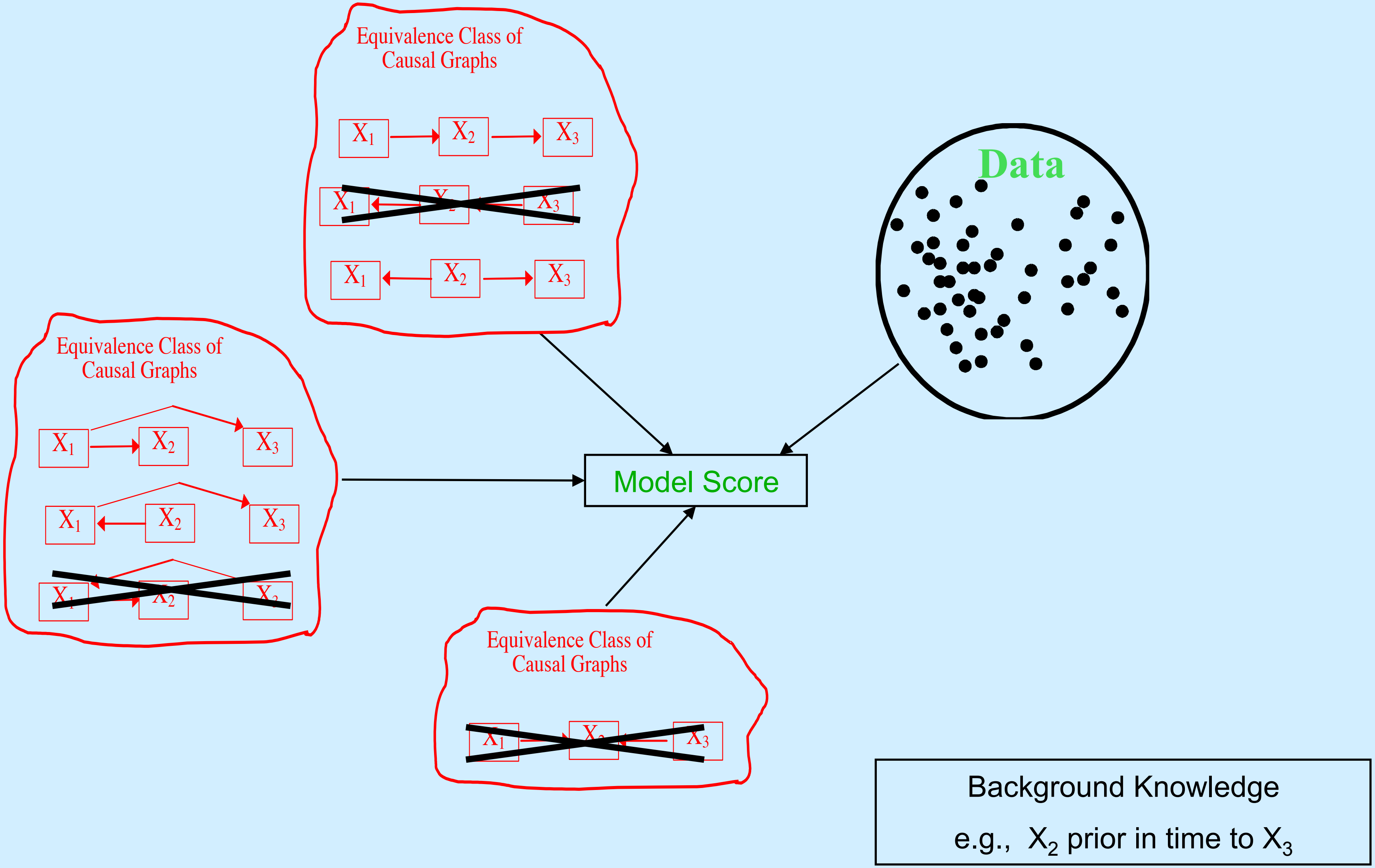
Constraint Based Search



X_1 _||_ X_2 means: $P(X_1, X_2) = P(X_1)P(X_2)$

X_1 _||_ $X_2 \mid X_3$ means: $P(X_1, X_2 \mid X_3) = P(X_1 \mid X_3)P(X_2 \mid X_3)$

Score Based Search

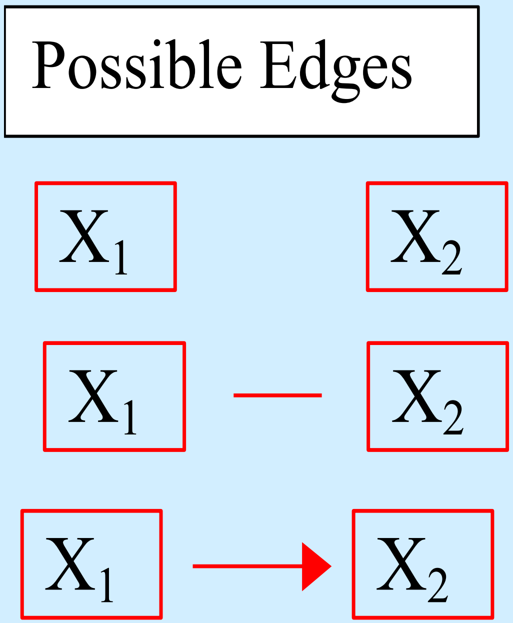


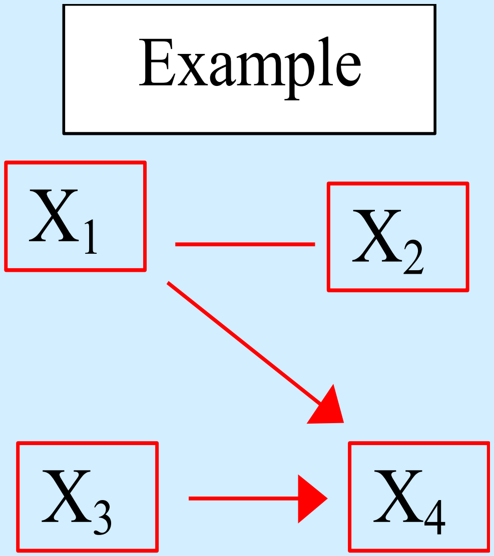
Independence Equivalence Classes: Patterns & PAGs

- Patterns (Verma and Pearl, 1990): graphical representation of d-separation equivalence among models with no latent common causes

- PAGs: (Richardson 1994) graphical representation of a d-separation equivalence class that includes models with *latent common causes* and *sample selection bias* that are d-separation equivalent over a set of measured variables **X**

Patterns





Patterns: What the Edges Mean

$\boxed{X_1} \qquad \boxed{X_2}$

X_1 and X_2 are not adjacent in any member of the equivalence class

---

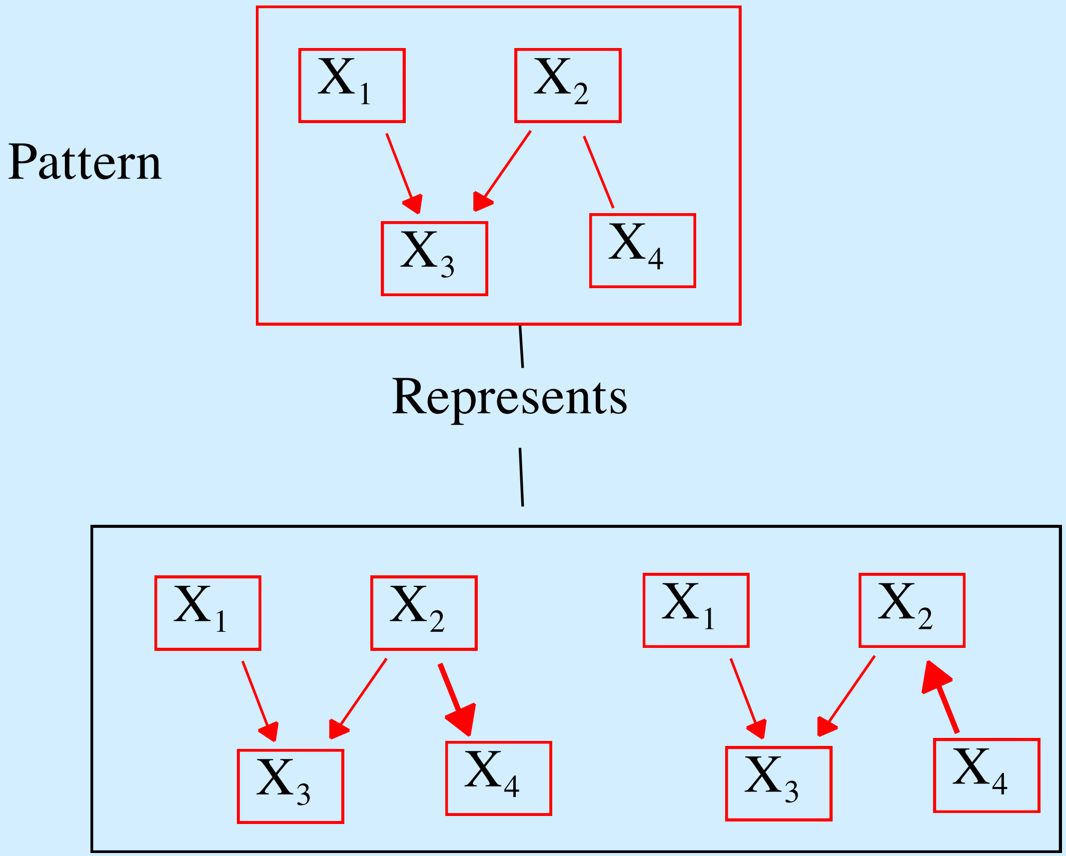
$\boxed{X_1} \longrightarrow \boxed{X_2}$

$X_1 \rightarrow X_2$ (X_1 is a cause of X_2) in every member of the equivalence class.

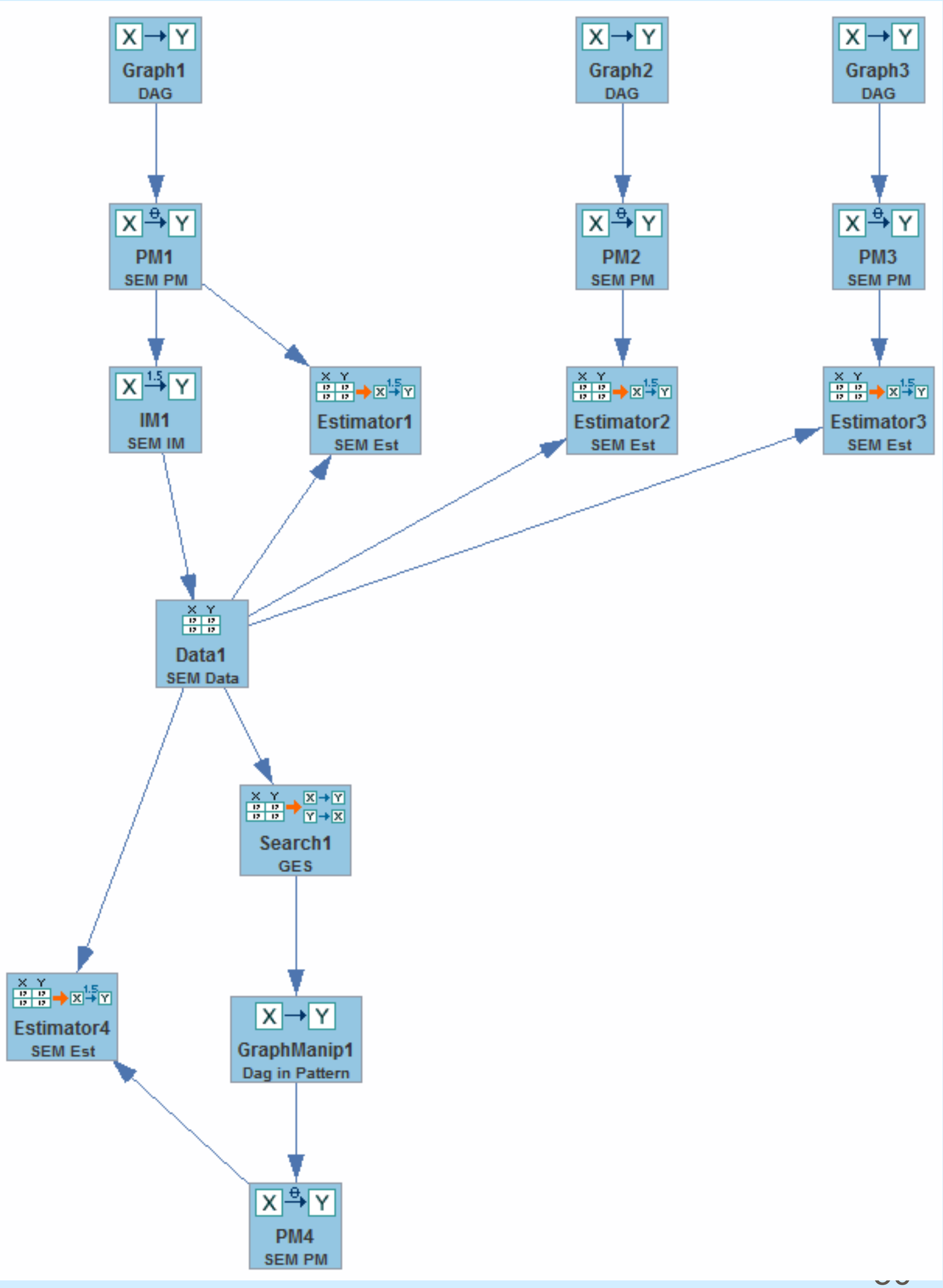
$\boxed{X_1} \; — \; \boxed{X_2}$

$X_1 \rightarrow X_2$ in some members of the equivalence class, and $X_2 \rightarrow X_1$ in others.

Patterns

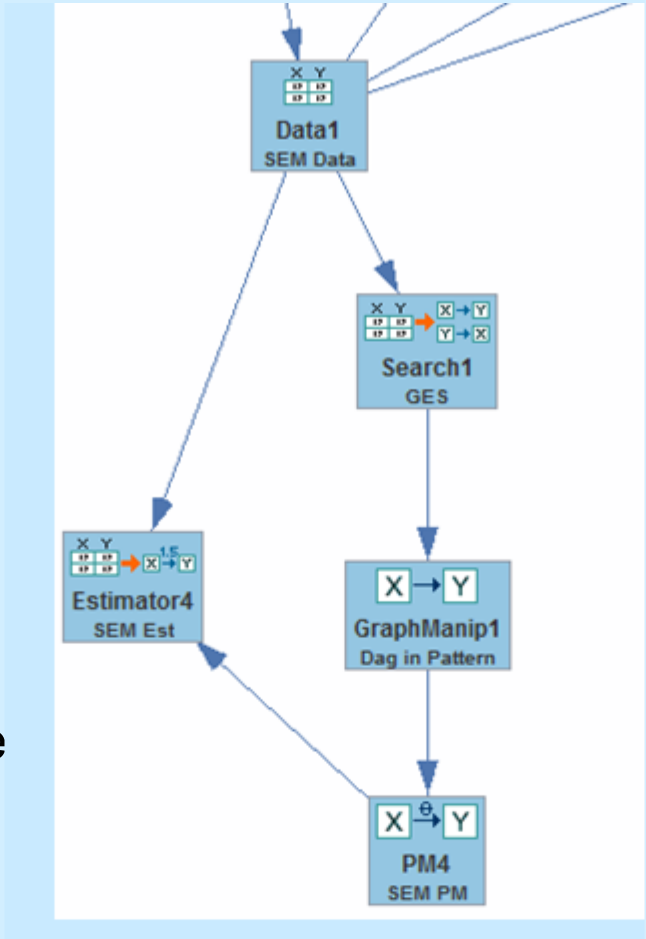


Tetrad Demo and Hands On



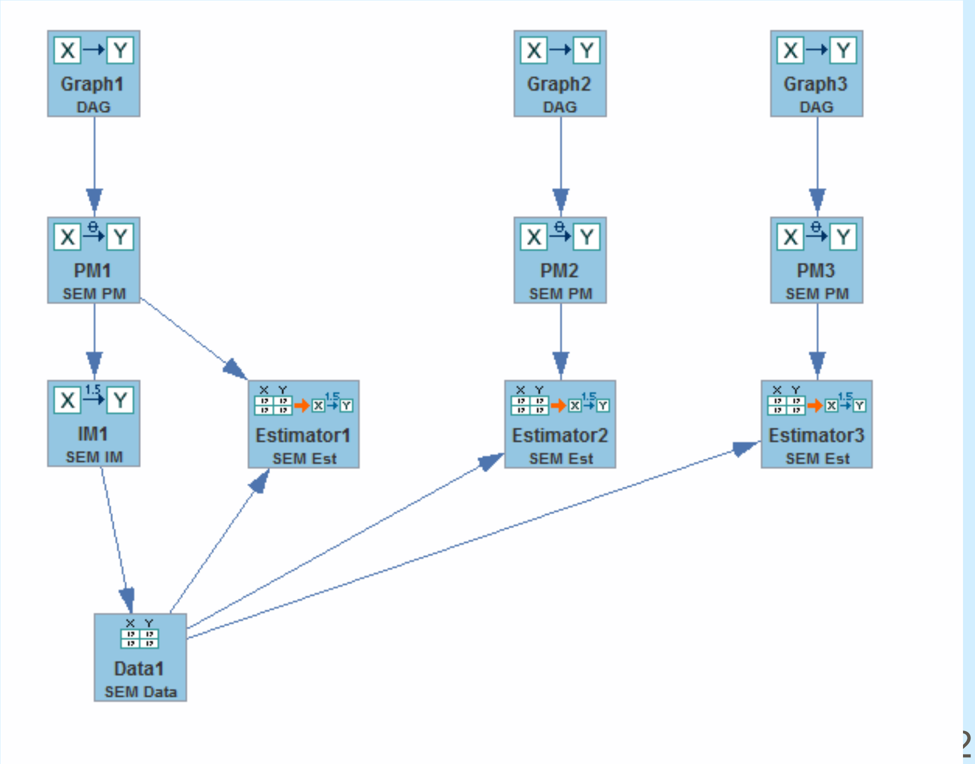
Tetrad Demo and Hands-on

1) Go to “session2”
2) Add Search node (from Data1)
   - Choose and execute one of the
   “Pattern searches”
3) Add a “Graph Manipulation” node to search result: “choose Dag in Pattern”
4) Add a PM to GraphManip
5) Estimate the PM on the data
6) Compare model-fit to model fit for true mode



Graphical Characterization of Model Equivalence

Why do some changes to the true model result in an equivalent model, but some do not?



d-separation/Independence Equivalence

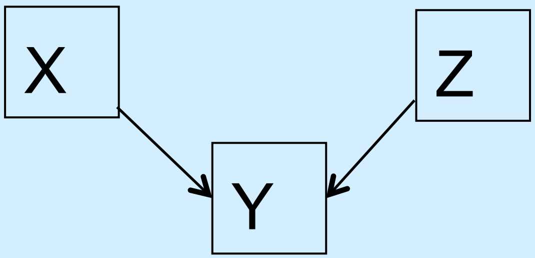
D-separation Equivalence Theorem (Verma and Pearl, 1988)

Two acyclic graphs over the same set of variables are d-separation equivalent iff they have:

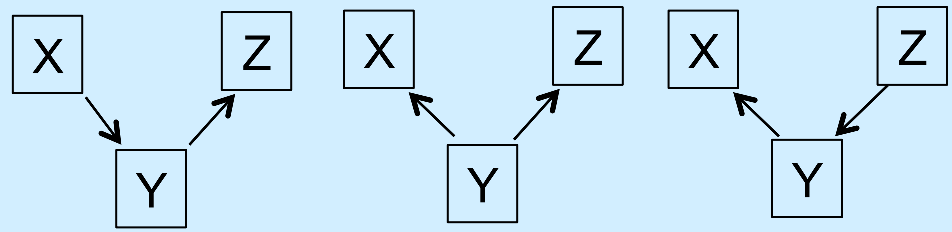
- the same adjacencies
- the same unshielded colliders

Colliders

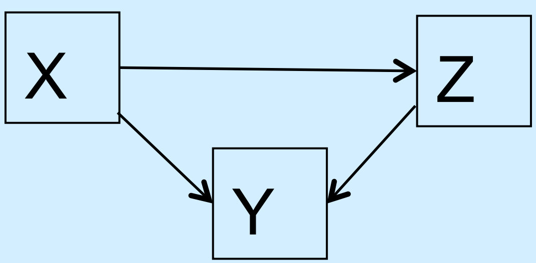
Y: *Collider*



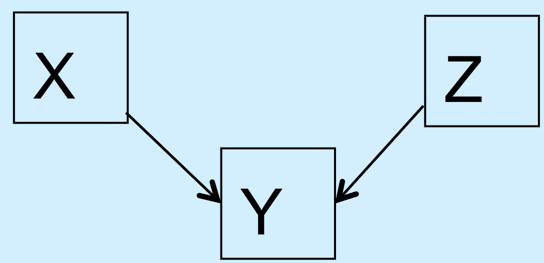
Y: *Non-Collider*



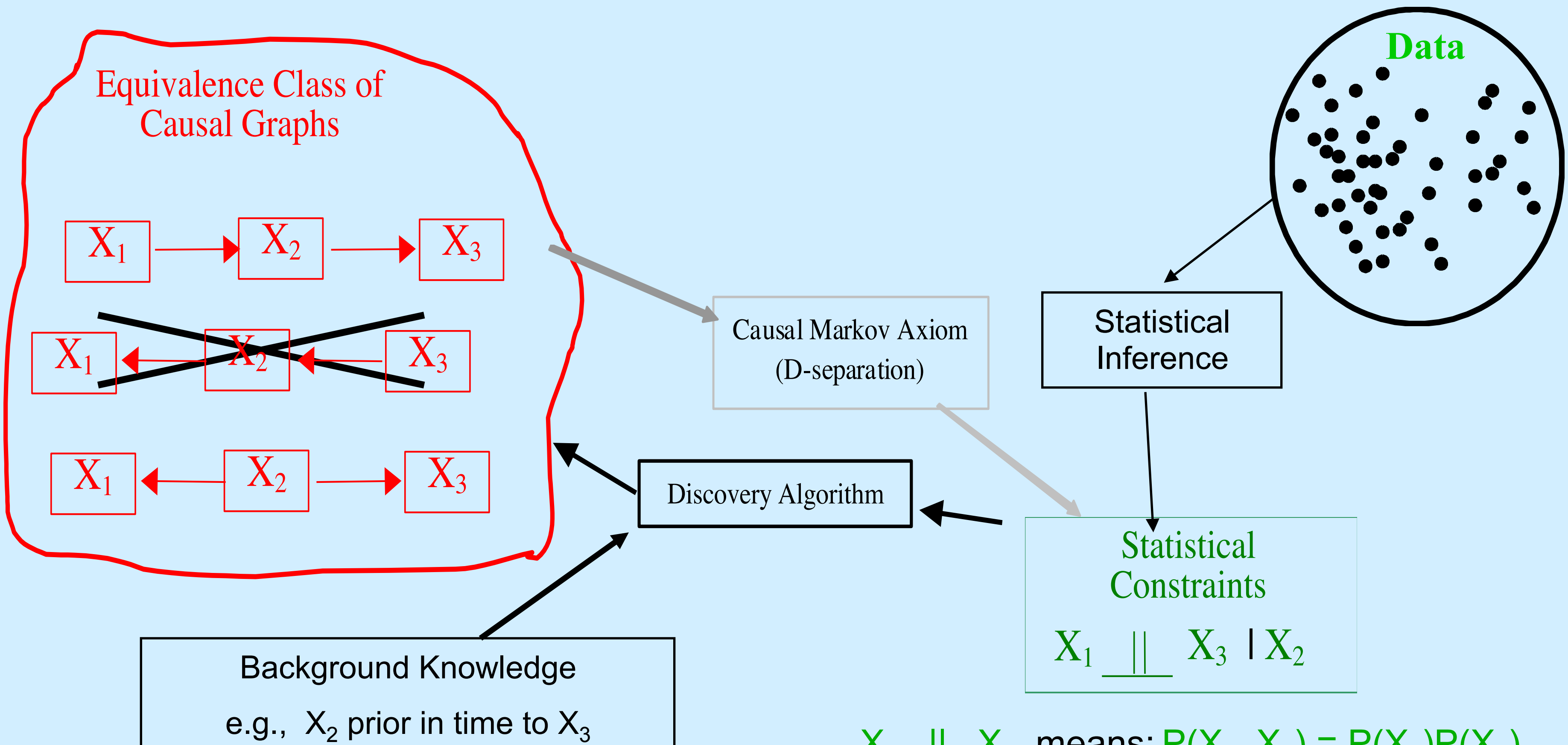
Shielded



Unshielded



Constraint Based Search

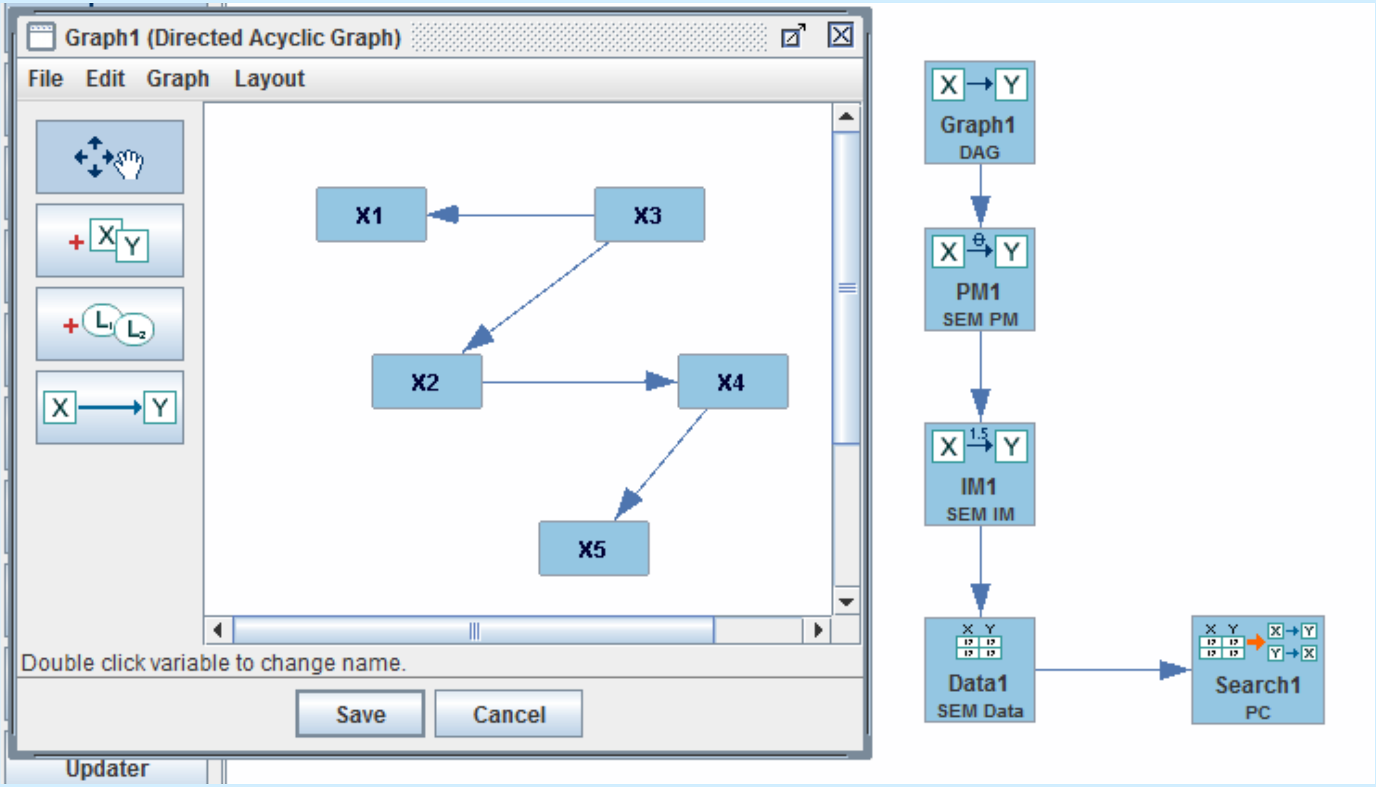


X_1 _||_ X_2 means: $P(X_1, X_2) = P(X_1)P(X_2)$

X_1 _||_ $X_2 \mid X_3$ means: $P(X_1, X_2 \mid X_3) = P(X_1 \mid X_3)P(X_2 \mid X_3)$

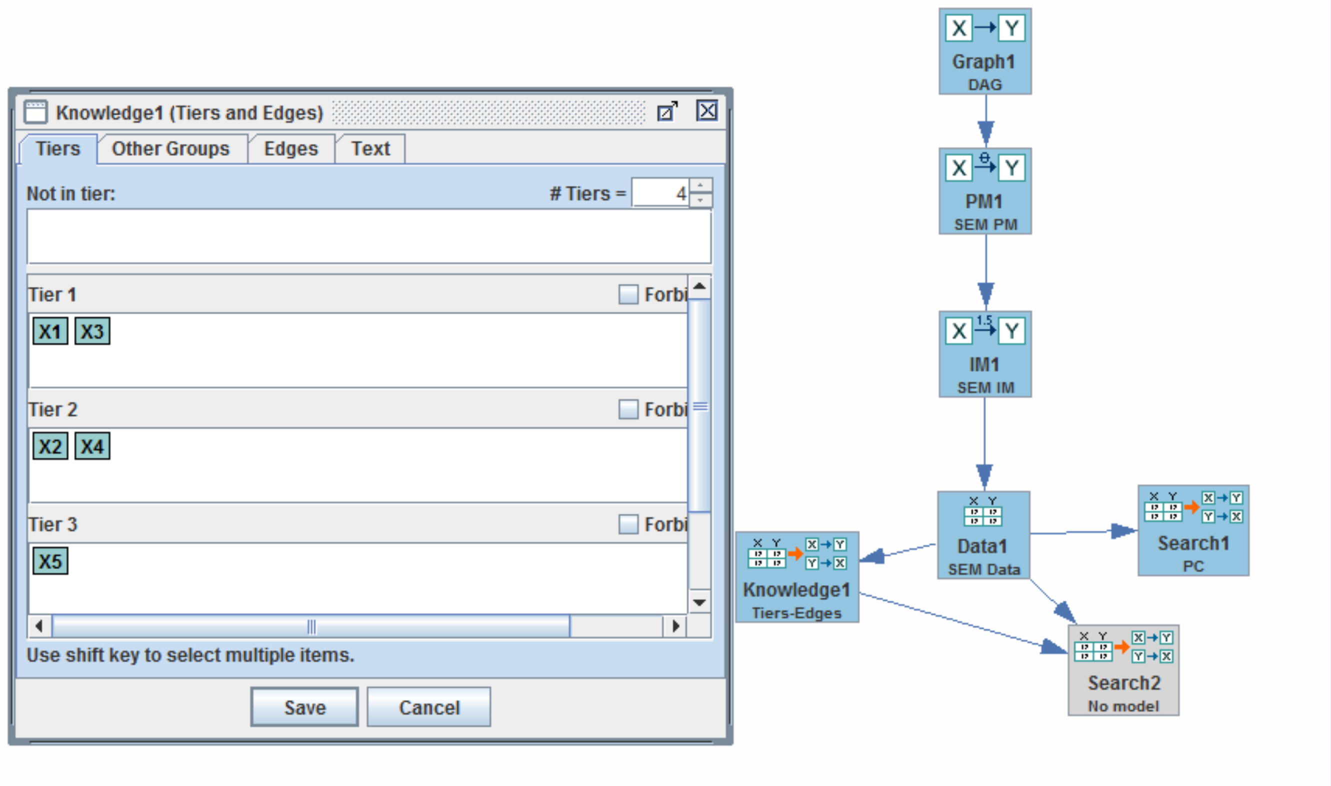
Backround Knowledge
Tetrad Demo and Hands-on

1) Create new session
2) Select "Search from Simulated Data" from Template menu
3) Build graph below, PM, IM, and generate sample data N=1,000.
4) Execute PC search, α = .05

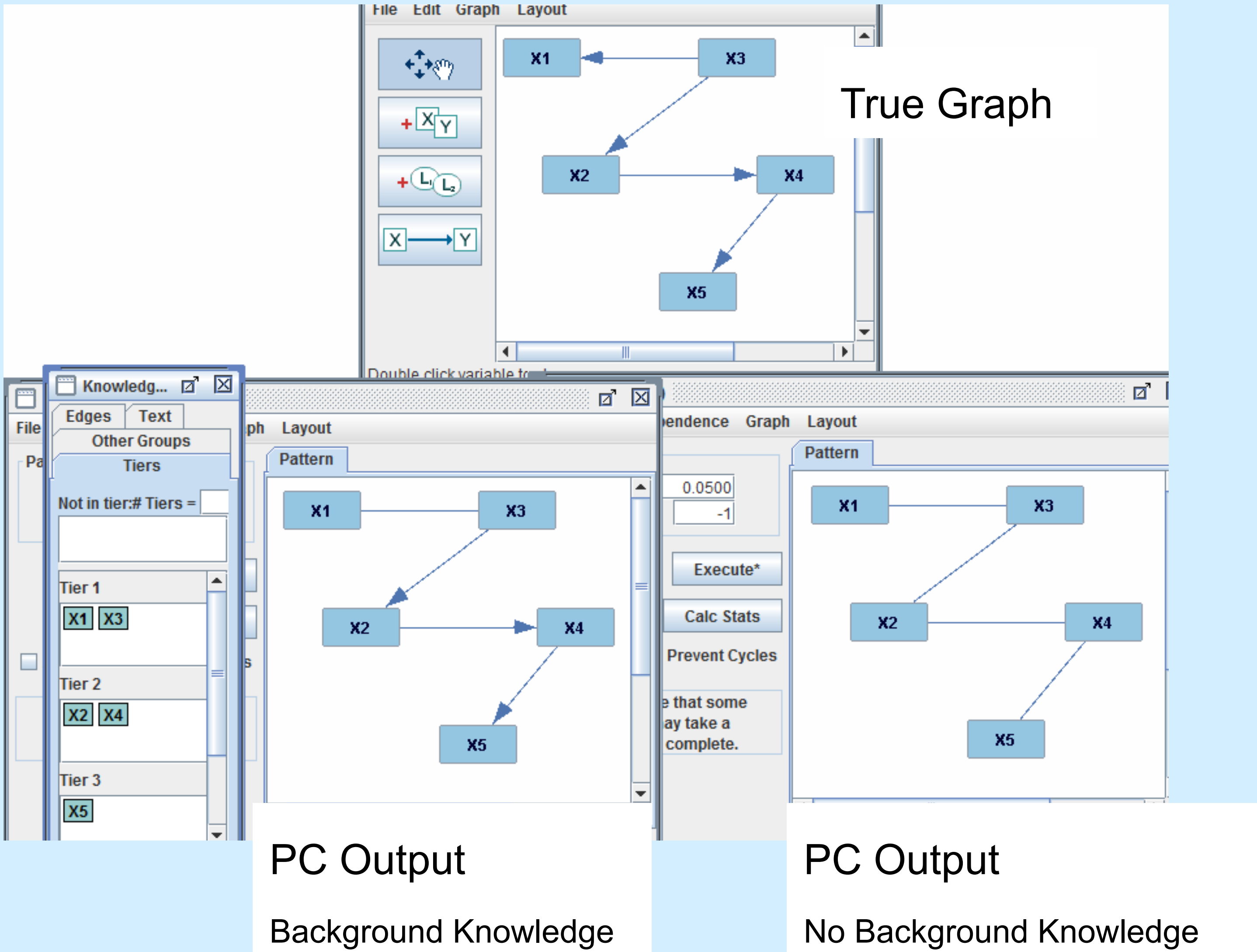


Backround Knowledge
Tetrad Demo and Hands-on

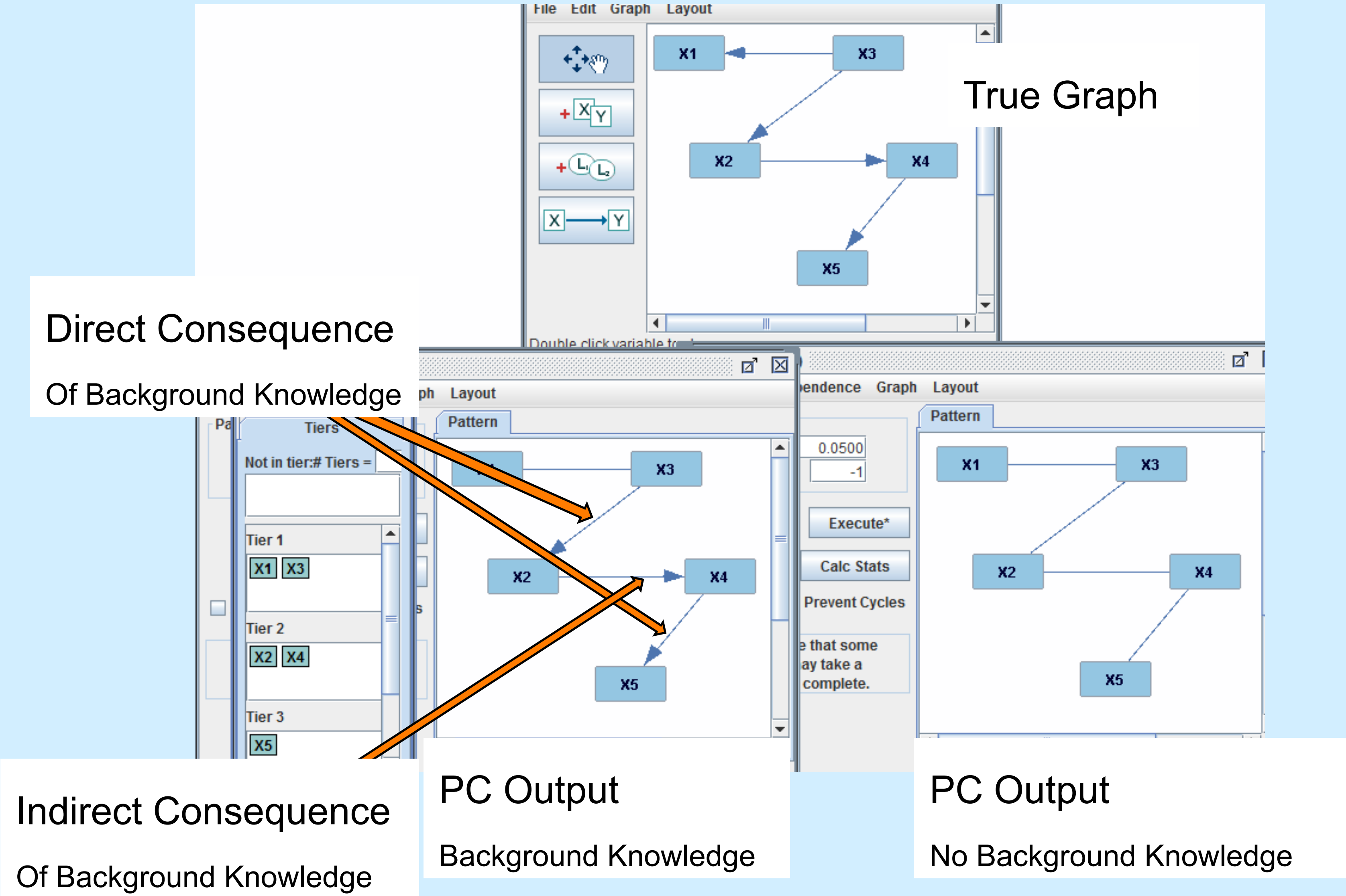
1) Add “Knowledge” node – as below
2) Create “Tiers” as shown below.
3) Execute PC search again, α = .05
4) Compare results (Search2) to previous search (Search1)



Backround Knowledge
Direct and Indirect Consequences



Backround Knowledge
Direct and Indirect Consequences



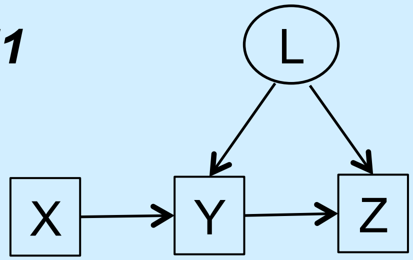
Independence Equivalence Classes: Patterns & PAGs

- Patterns (Verma and Pearl, 1990): graphical representation of d-separation equivalence among models with no latent common causes

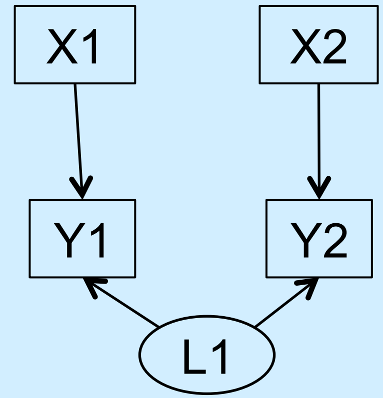
- PAGs: (Richardson 1994) graphical representation of a d-separation equivalence class that includes models with *latent common causes* and *sample selection bias* that are d-separation equivalent over a set of measured variables **X**

Interesting Cases

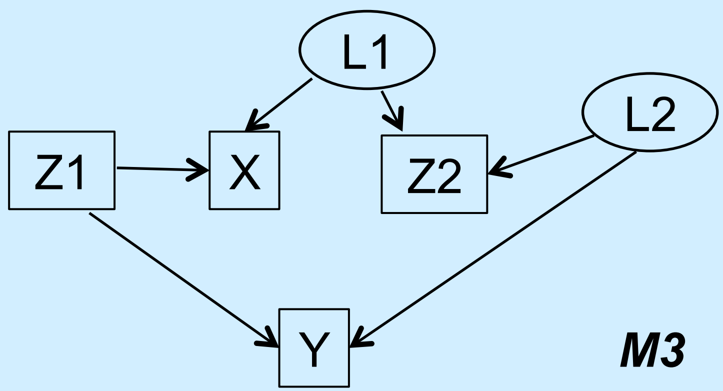
M1



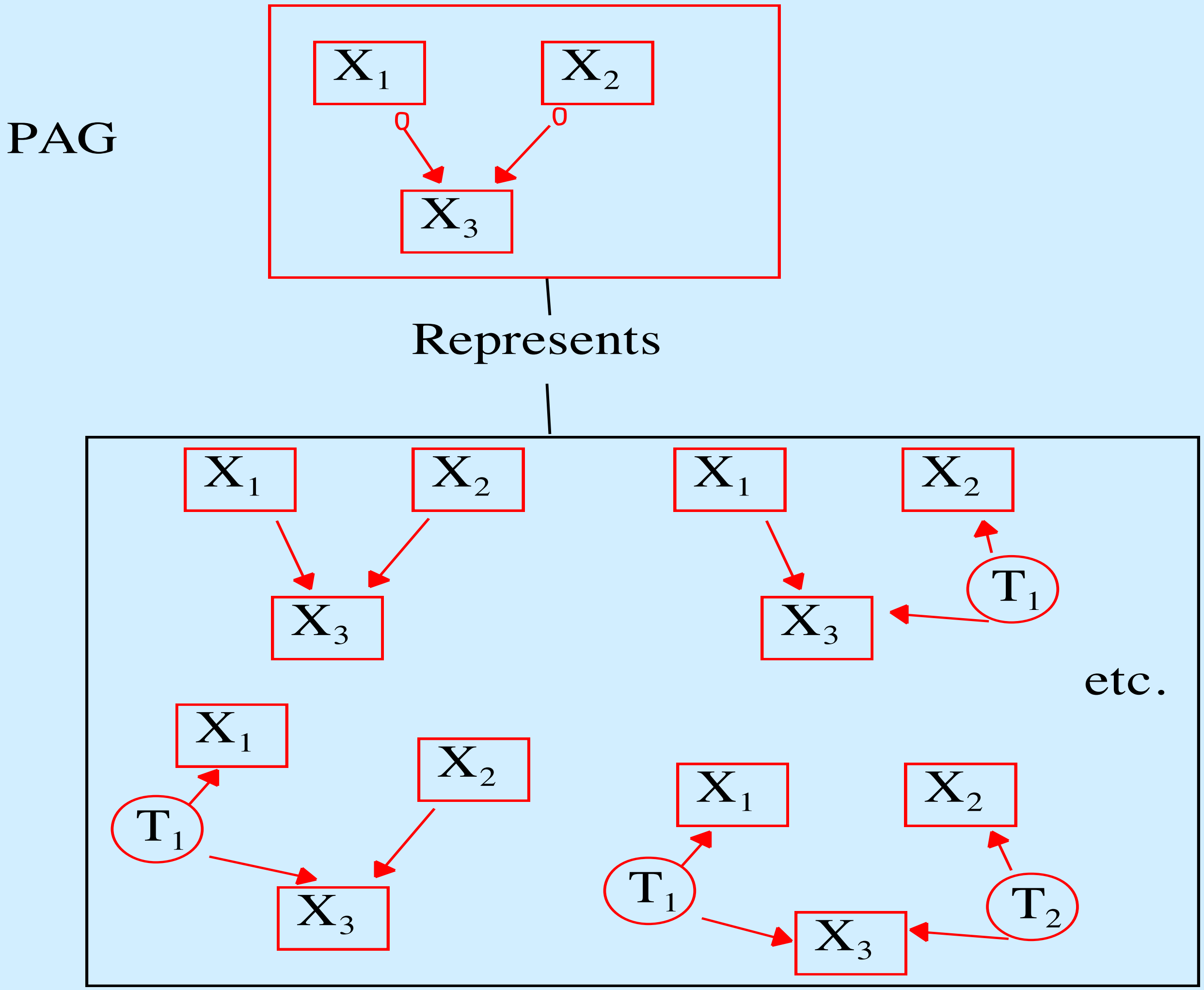
M2



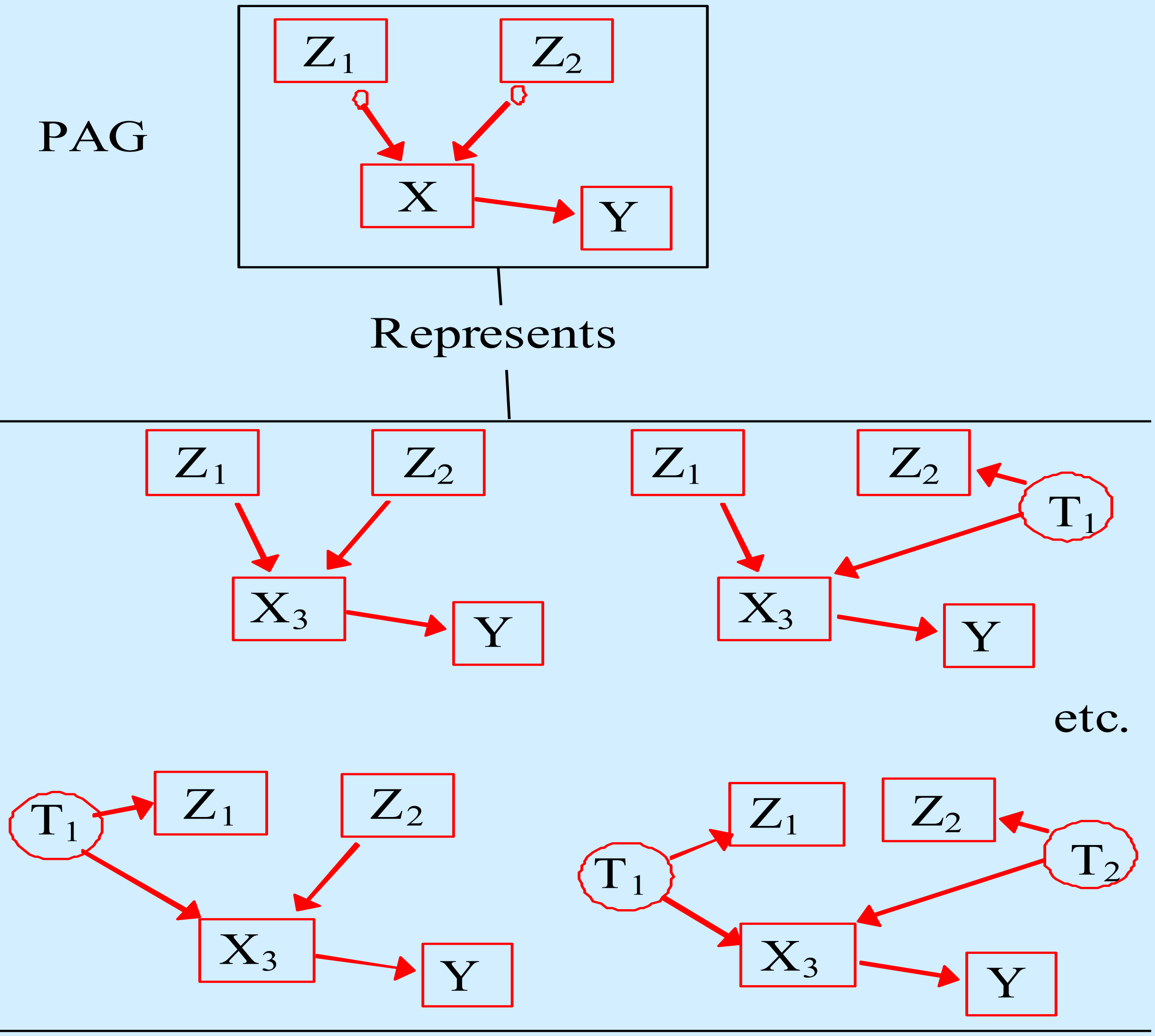
M3



PAGs: Partial Ancestral Graphs



PAGs: Partial Ancestral Graphs



PAGs: Partial Ancestral Graphs

What PAG edges mean.

$X_1 \quad X_2$	X_1 and X_2 are not adjacent
$X_1 \circ\!\!\rightarrow X_2$	X_2 is not an ancestor of X_1
$X_1 \circ\!\!-\!\!\circ X_2$	No set d-separates X_2 and X_1
$X_1 \rightarrow X_2$	X_1 is a cause of X_2
$X_1 \leftrightarrow X_2$	There is a latent common cause of X_1 and X_2

Tetrad Demo and Hands-on

1) Create new session

2) Select “Search from Simulated Data” from Template menu

3) Build graph below, SEM PM, IM, and generate sample data N=1,000.

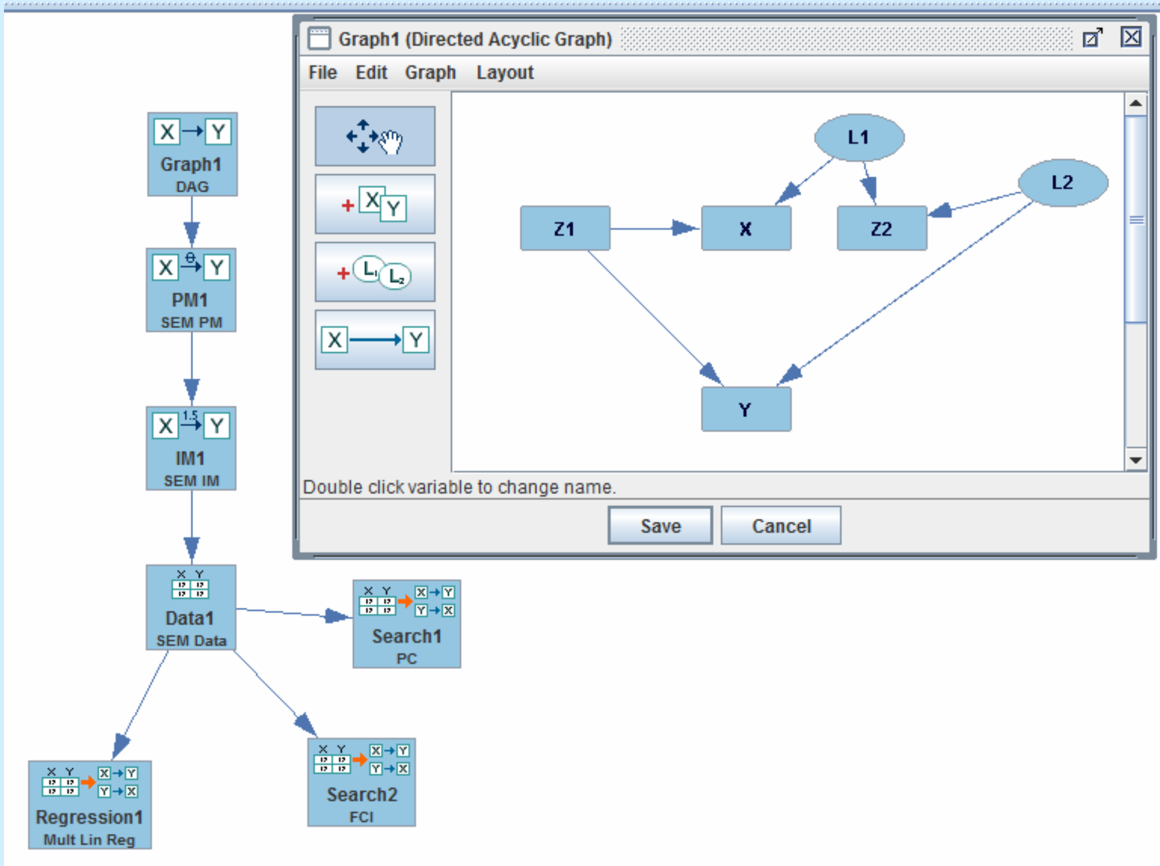
4) Execute PC search, α = .05

5) Execute FCI search, α = .05

6) Estimate multiple regression,

   Y as response,

   Z1, X, Z2 as Predictors



Search Methods

- Constraint Based Searches
  - PC, FCI
  - Very fast – capable of handling >5,000 variables
  - Pointwise, but not uniformly consistent

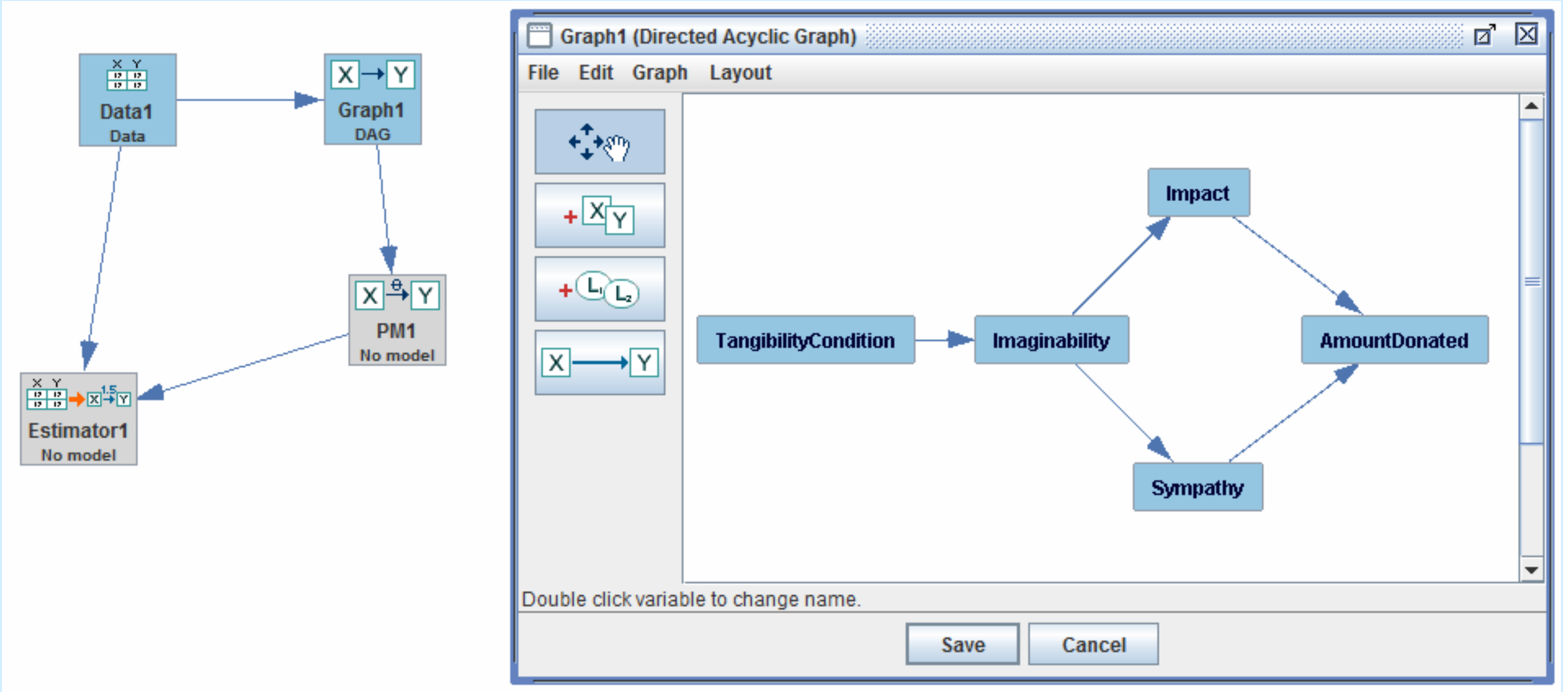
- Scoring Searches
  - Scores: BIC, AIC, etc.
  - Search: Hill Climb, Genetic Alg., Simulated Annealing
  - Difficult to extend to latent variable models
  - Meek and Chickering Greedy Equivalence Class (GES)
  - Slower than constraint based searches – but now capable of 1,000 vars
  - Pointwise, but not uniformly consistent

- Latent Variable Psychometric Model Search
  - BPC, MIMbuild, etc.

- Linear non-Gaussian models (Lingam)
- Models with cycles
- And more!!!

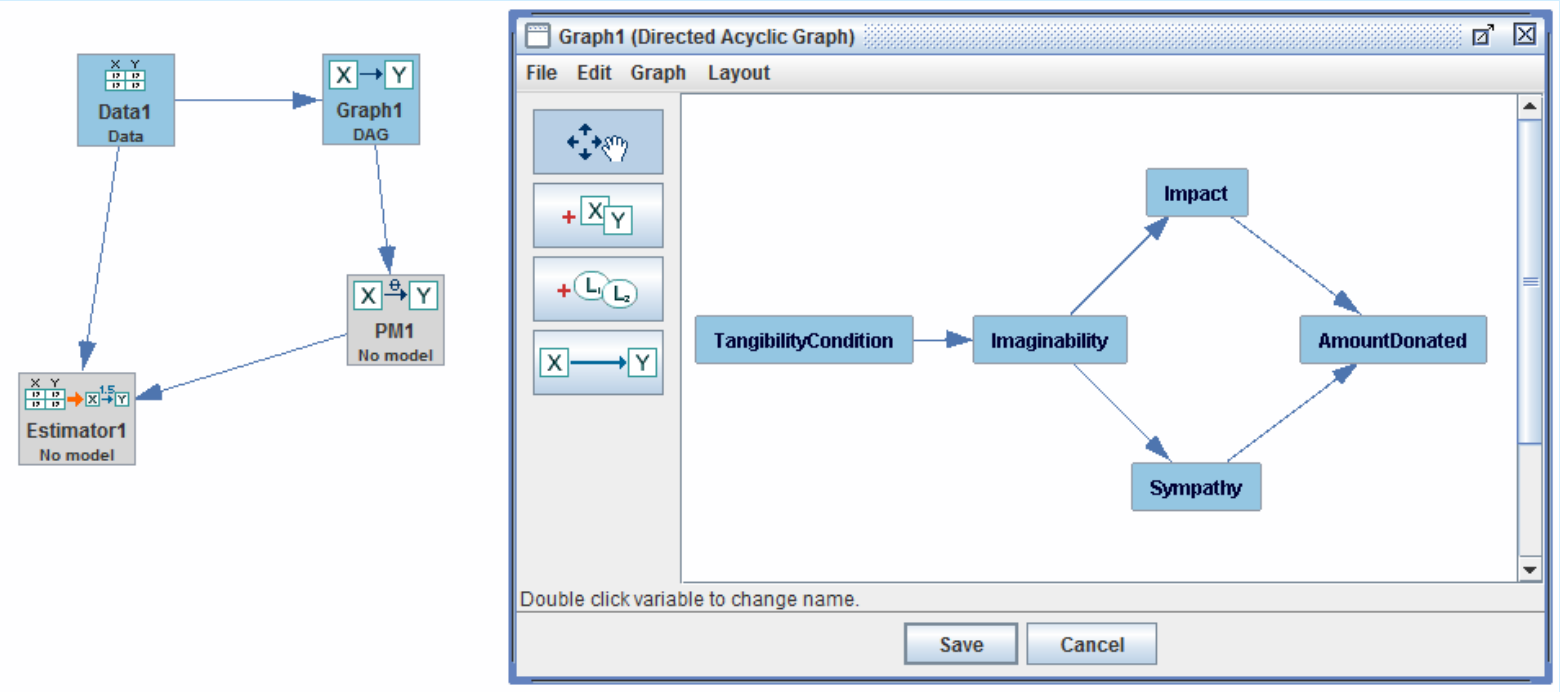
Tetrad Demo and Hands-on

1) Load charity.txt (tabular – not covariance data)
2) Build graph of theoretical hypothesis
3) Build SEM PM from graph
4) Estimate PM, check results



Tetrad Demo and Hands-on

1) Create background knowledge: Tangibility exogenous (uncaused)
2) Search for models
3) Estimate one model from the output of search
4) Check model fit, check parameter estimates, esp. their sign



Thank You!

Additional Slides

Constraint-based Search

1) Adjacency
2) Orientation

Constraint-based Search: Adjacency

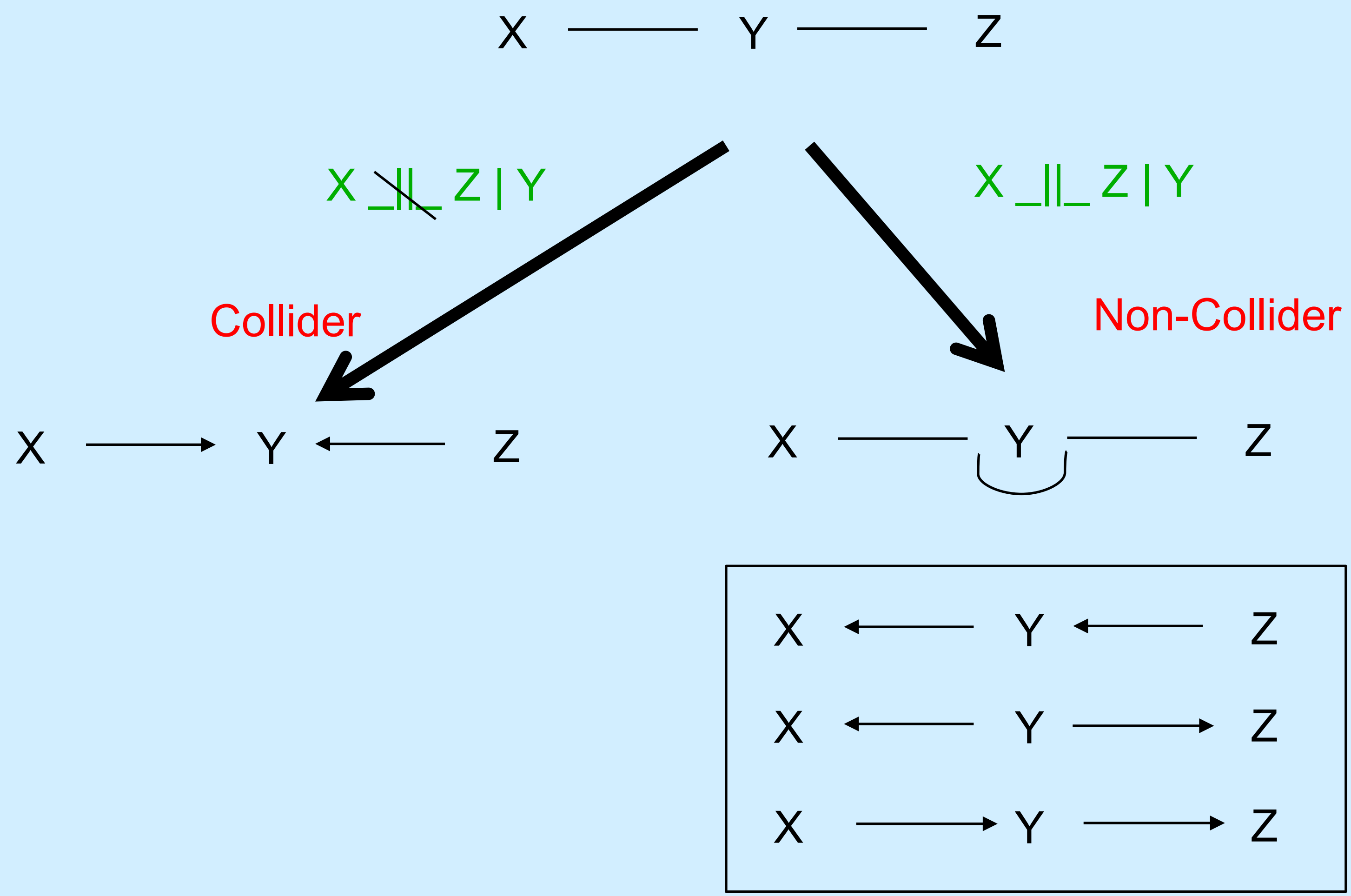
1. *X and Y are <u>adjacent</u> if they are dependent conditional on all subsets that don't include them*

2. *X and Y are <u>not adjacent</u> if they are independent conditional on <u>any</u> subset that doesn't include them*

Search: Orientation

Patterns

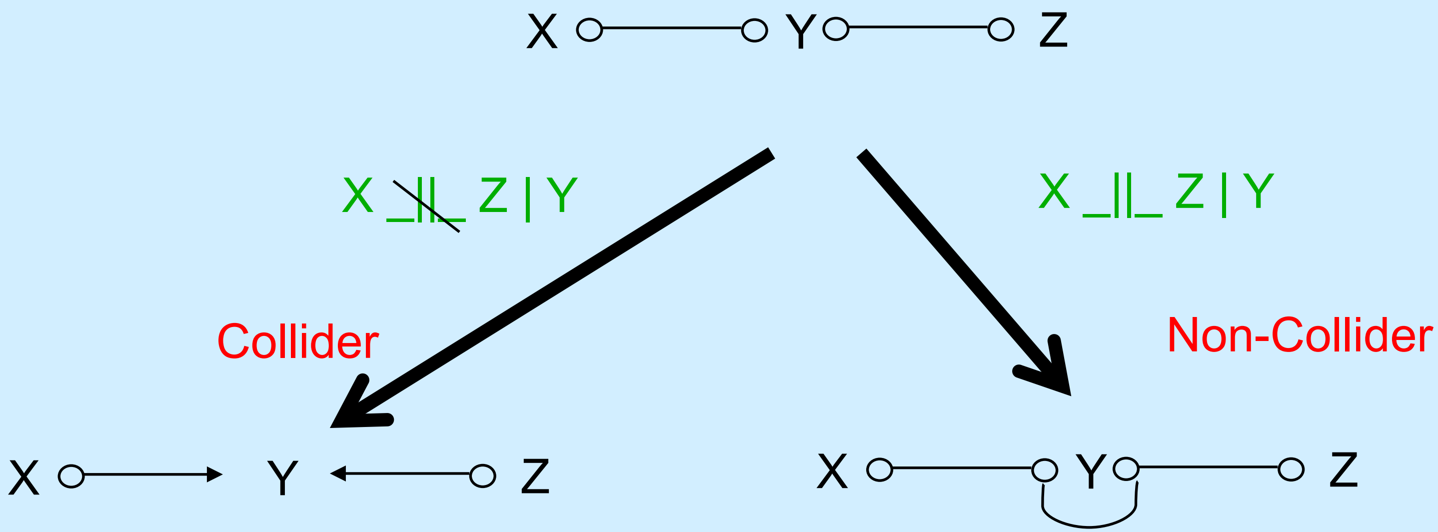
Y Unshielded



Search: Orientation

PAGs

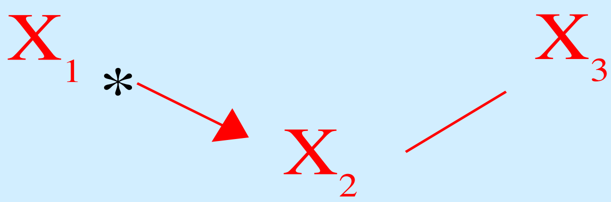
Y Unshielded



Search: Orientation

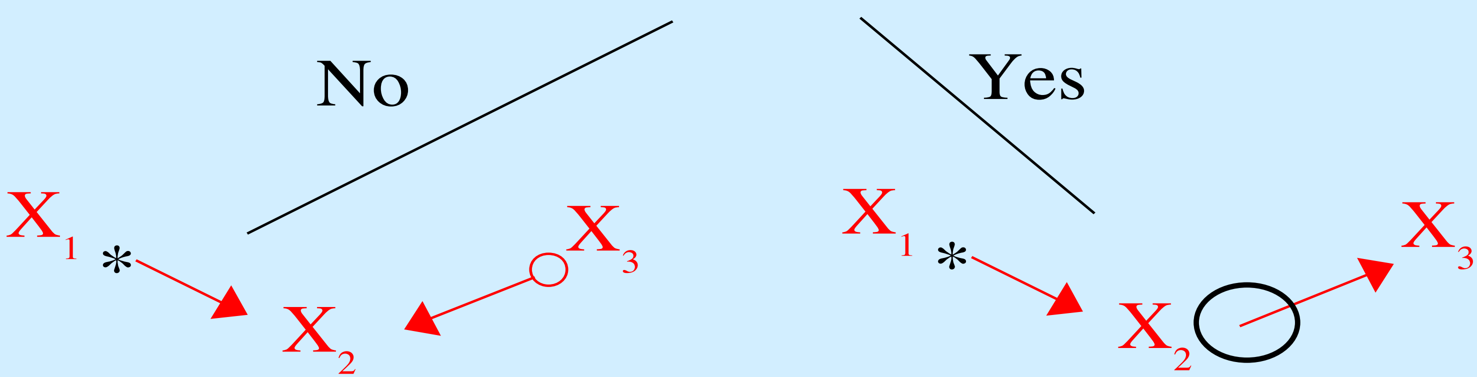
Away from Collider

Test Conditions

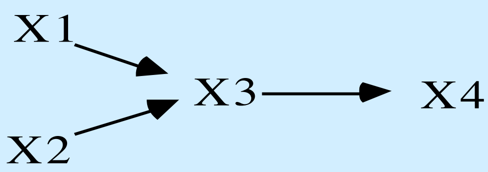


1) X_1 - X_2 adjacent, and *into* X_2.
2) X_2 - X_3 adjacent
3) X_1 - X_3 not adjacent

Test $X_1 \perp\!\!\!\perp X_3 \mid X_2$



Causal Graph



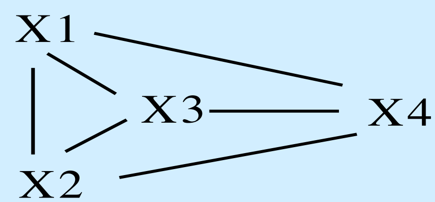
Independcies

$X1 \perp\!\!\!\perp X2$

$X1 \perp\!\!\!\perp X4 \mid \{X3\}$

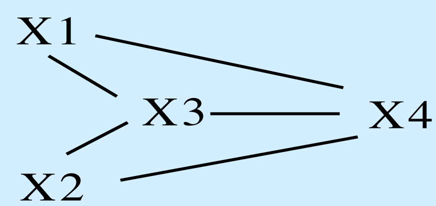
$X2 \perp\!\!\!\perp X4 \mid \{X3\}$

Begin with:



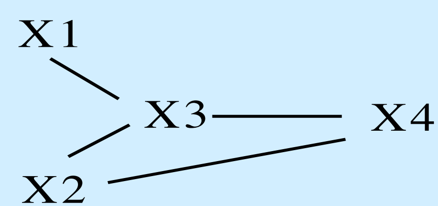
From

$X1 \perp\!\!\!\perp X2$



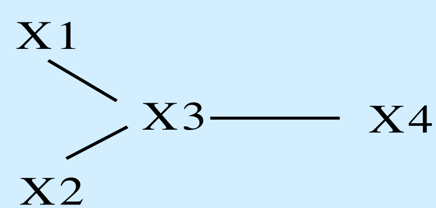
From

$X1 \perp\!\!\!\perp X4 \mid \{X3\}$



From

$X2 \perp\!\!\!\perp X4 \mid \{X3\}$



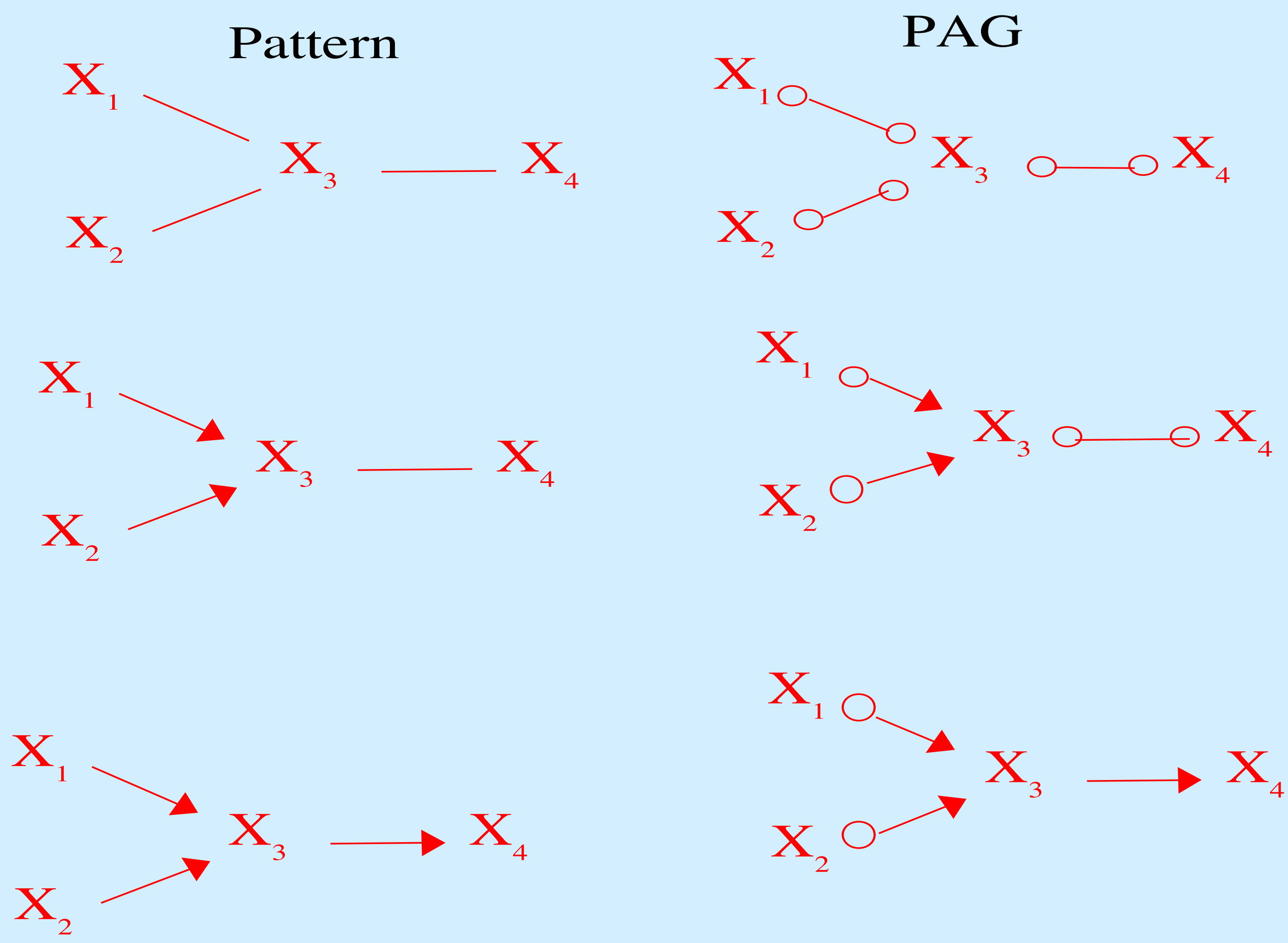
Search: Orientation

After Orientation Phase

$X1 \perp\!\!\!\perp X2$

$X1 \perp\!\!\!\perp X4 \mid X3$

$X2 \perp\!\!\!\perp X4 \mid X3$



Bridge Principles:
Acyclic Causal Graph over V $\Rightarrow$ Constraints on P(V)

Weak Causal Markov Assumption

V_1,V_2 causally disconnected $\Rightarrow$ V_1 _||_ V_2

V_1,V_2 causally disconnected $\Leftrightarrow$

i. V_1 not a cause of V_2, and

ii. V_2 not a cause of V_1, and

iii. No common cause Z of V_1 and V_2

V_1 _||_ V_2 $\Leftrightarrow$ $P(V_1,V_2) = P(V_1)P(V_2)$

Bridge Principles:

Acyclic Causal Graph over V ⇒ Constraints on P(V)

Weak Causal Markov Assumption

V_1,V_2 causally disconnected ⇒ V_1 _||_ V_2

Determinism

(Structural Equations)

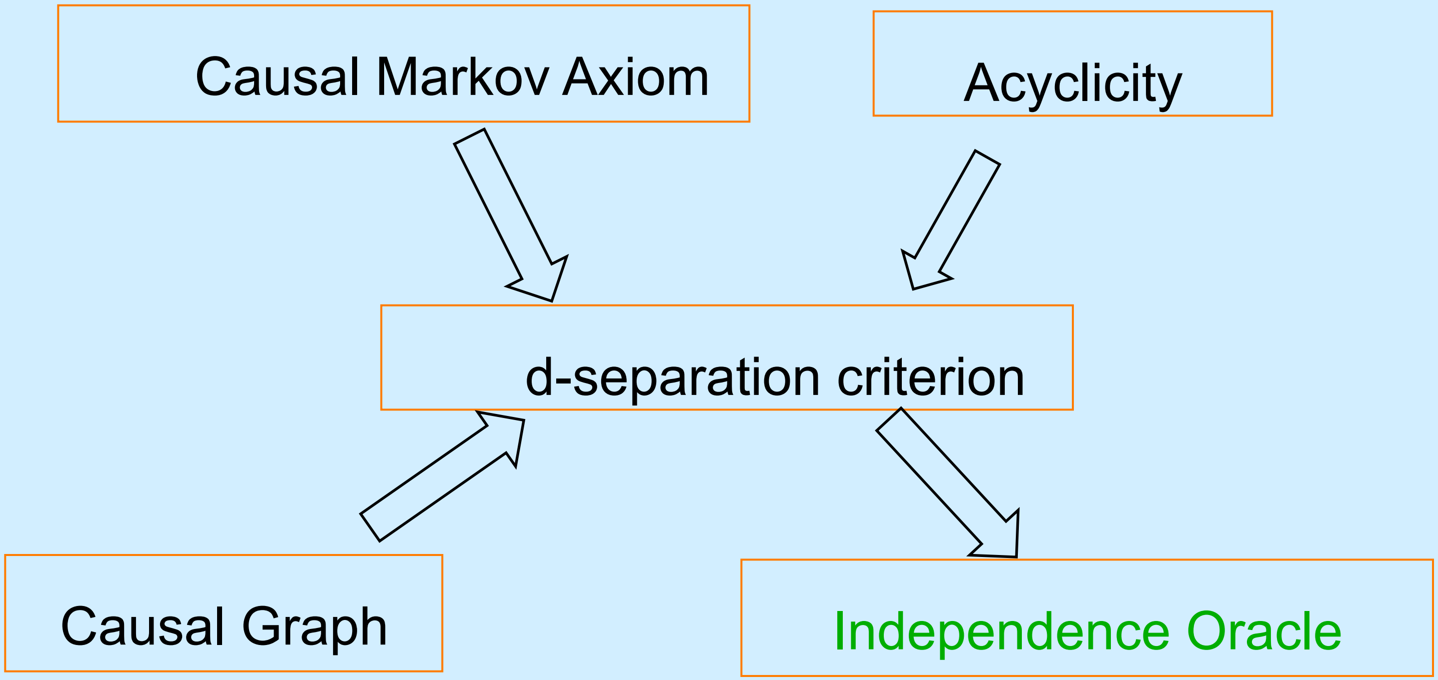
Causal Markov Axiom

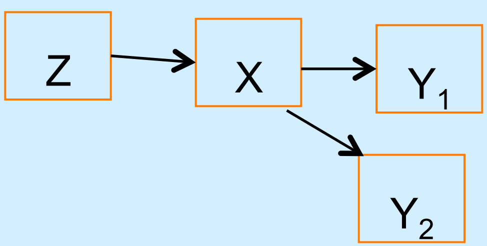
If G is a causal graph, and P a probability distribution over the variables in G, then in <G,P> satisfy the Markov Axiom iff:

every variable V is independent of its non-effects,

conditional on its immediate causes.

Bridge Principles:
Acyclic Causal Graph over V ⇒ Constraints on P(V)

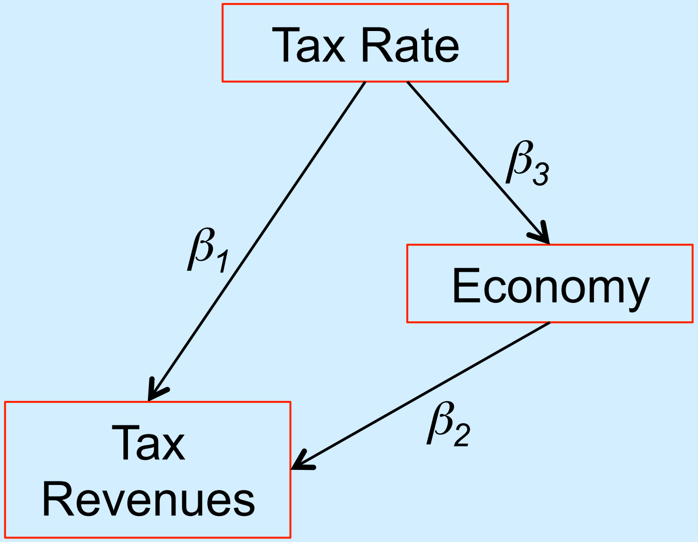




$Z \perp\!\!\!\perp Y_1 \mid X$	$Z \perp\!\!\!\perp Y_2 \mid X$
$Z \perp\!\!\!\perp Y_1 \mid X, Y_2$	$Z \perp\!\!\!\perp Y_2 \mid X, Y_1$
$Y_1 \perp\!\!\!\perp Y_2 \mid X$	$Y_1 \perp\!\!\!\perp Y_2 \mid X, Z$

Faithfulness

Constraints on a probability distribution P generated by a causal structure G hold for all parameterizations of G.



Revenues := β_1*Rate* + β_2*Economy* + ε_{Rev}

Economy := β_3*Rate* + ε_{Econ}

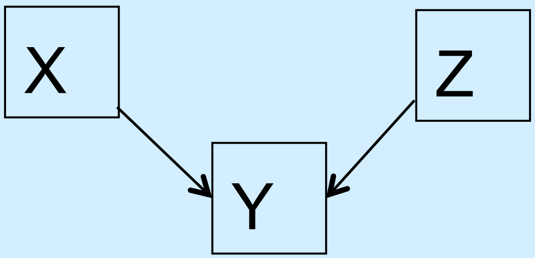
Faithfulness:

$\beta_1 \neq -\beta_3\beta_2$

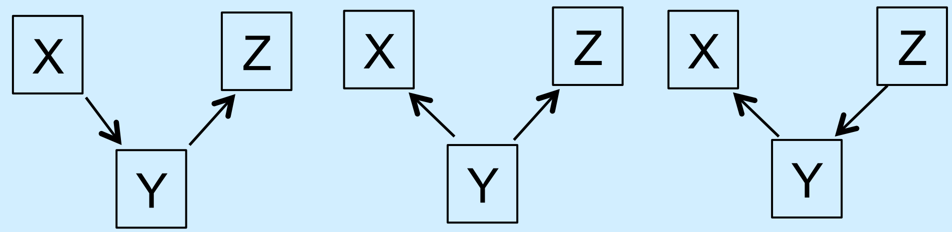
$\beta_2 \neq -\beta_3\beta_1$

Colliders

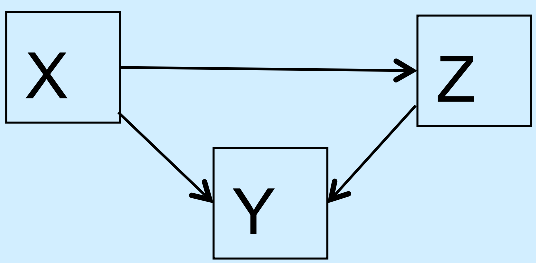
Y: *Collider*



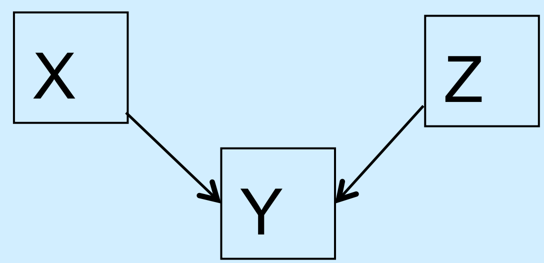
Y: *Non-Collider*



Shielded

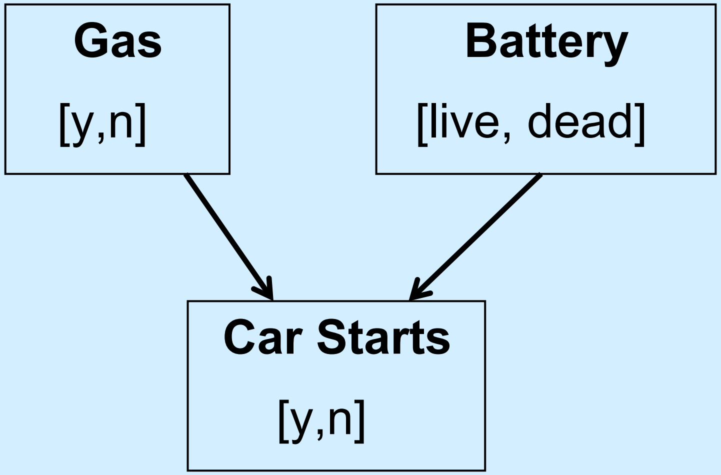


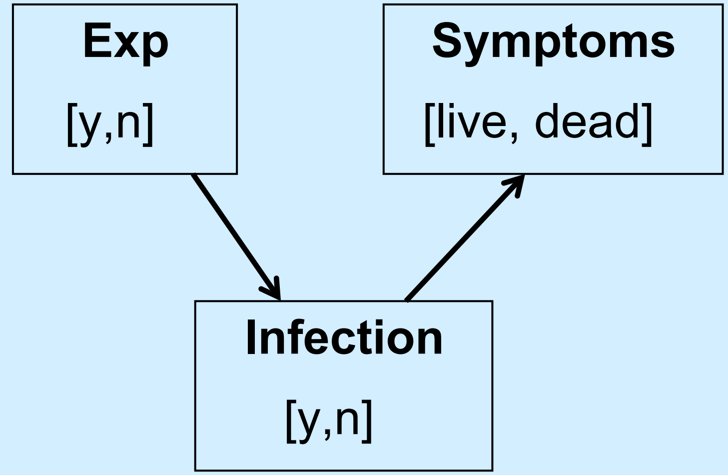
Unshielded



Colliders induce Association

Non-Colliders screen-off Association





Gas _||_ Battery

Gas ~~_||_~~ Battery | Car starts = no

Exp ~~_||_~~ Symptoms

Exp _||_ Symptoms | Infection

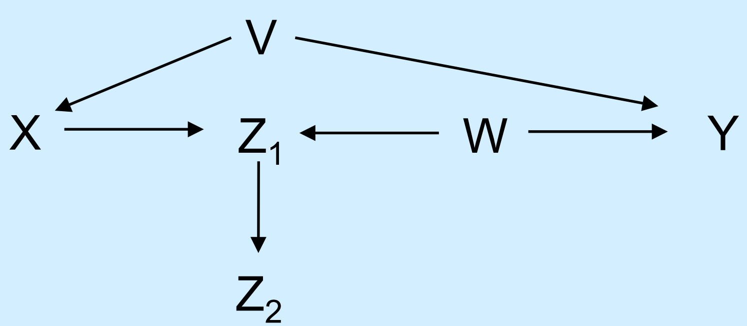
D-separation

X is *d-separated* from Y by **Z** in **G** iff
Every undirected path between X and Y in **G** is *inactive* relative to **Z**

An undirected path is *inactive* relative to **Z** iff
any node on the path is *inactive* relative to **Z**

A node N (on a path) is *inactive* relative to **Z** iff
a) N is a non-collider in Z, or
b) N is a collider that is not in Z, and has no descendant in Z

A node N (on a path) is *active* relative to **Z** iff
a) N is a non-collider not in Z, or
b) N is a collider that is in Z, or has a descendant in Z



Undirected Paths between X , Y:

1) X --> Z_1 <-- W --> Y

2) X <-- V --> Y

D-separation

X is *d-separated* from Y by **Z** in **G** iff
Every undirected path between X and Y in **G** is inactive relative to **Z**

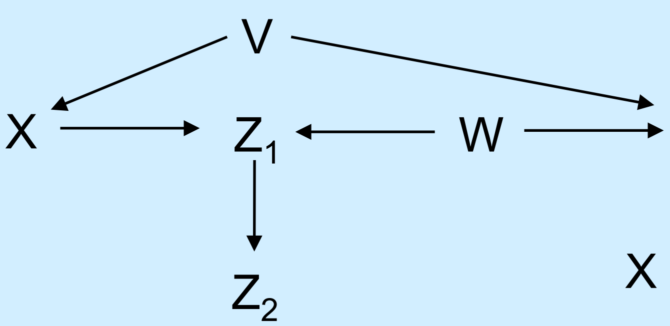
An undirected path is inactive relative to **Z** iff
any node on the path is inactive relative to **Z**

A node N is inactive relative to **Z** iff
a) N is a non-collider in Z, or
b) N is a collider that is not in Z,
and has no descendant in Z

Undirected Paths between X , Y:

1) X --> Z_1 <-- W --> Y

2) X <-- V --> Y



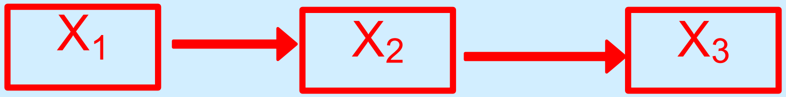
X d-sep Y relative to **Z** = ∅ ? *No*

X d-sep Y relative to **Z** = {V} ? *Yes*

X d-sep Y relative to **Z** = {V, Z_1 } ? *No*

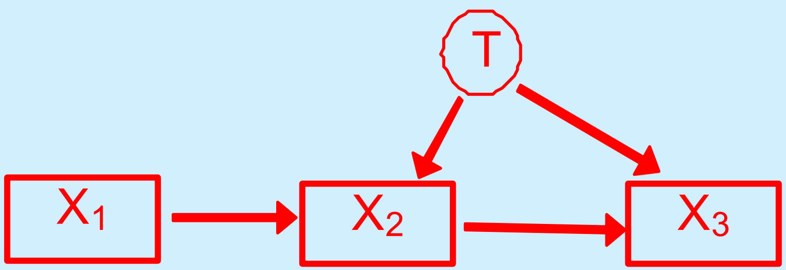
X d-sep Y relative to **Z** = {W, Z_2 } ? *Yes*

D-separation



X_3 and X_1 d-sep by X_2?

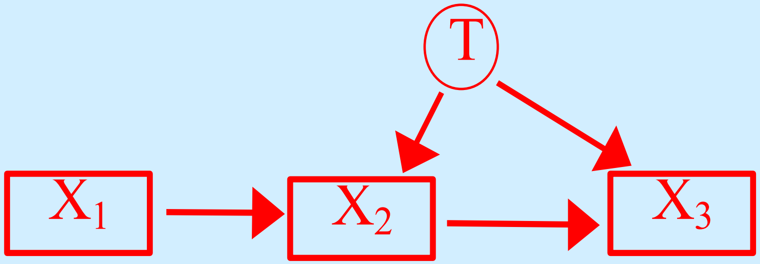
Yes: $X_3 \perp\!\!\!\perp X_1 \mid X_2$



X_3 and X_1 d-sep by X_2?

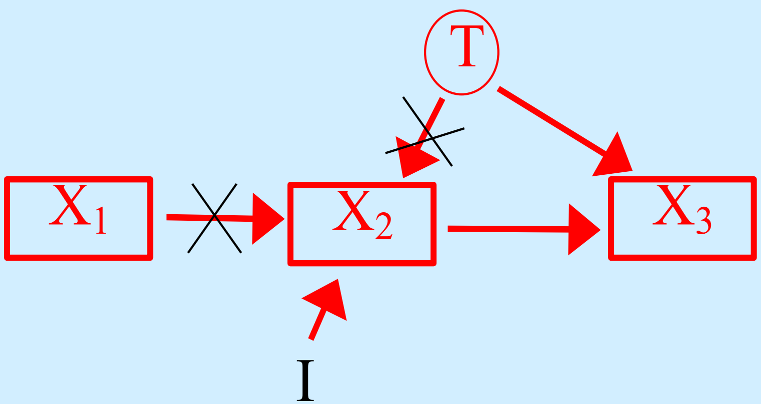
No: $X_3 \not\perp\!\!\!\perp X_1 \mid X_2$

Statistical Control ≠ Experimental Control



Statistically control for X_2

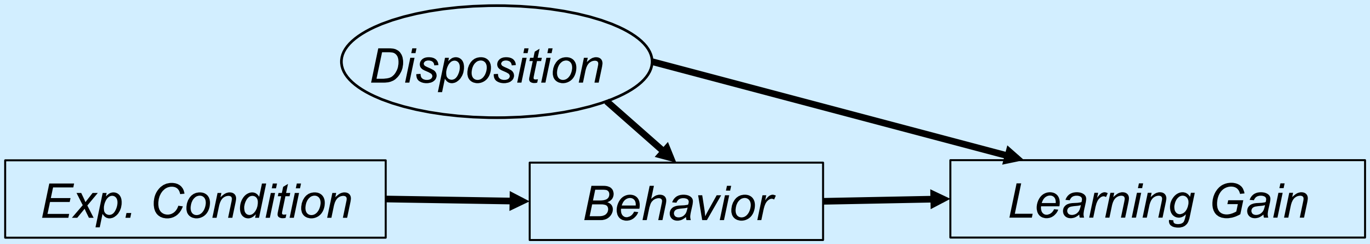
$X_3 \not\perp\!\!\!\perp X_1 \mid X_2$



Experimentally control for X_2

$X_3 \perp\!\!\!\perp X_1 \mid X_2(\text{set})$

Statistical Control ≠ Experimental Control



Exp. Cond ~~_||_~~ Learning Gain

Exp → Learning

Exp. Cond _||_ Learning Gain | Behavior, Disposition

Exp → Learning is
Mediated by Behavior

Exp. Cond _||_ Learning Gain | Behavior set

Exp → Learning is
Mediated by Behavior

Exp. Cond ~~_||_~~ Learning Gain | Behavior observed

Exp → Learning is *not*
Mediated by Behavior
or
Unmeasured Confounder

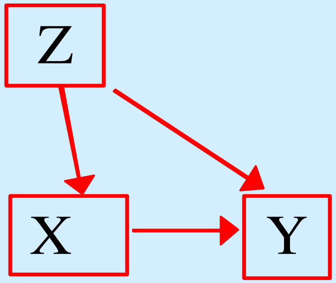
Regression
&
Causal Inference

Regression & Causal Inference

Typical (non-experimental) strategy:

1. Establish a prima facie case (X associated with Y)

But, omitted variable bias



2. So, identifiy and measure potential confounders **Z:**
   a) prior to X,
   b) associated with X,
   c) associated with Y

3. Statistically adjust for **Z** (multiple regression)

Regression & Causal Inference

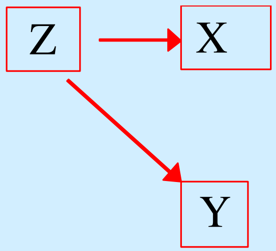
Strategy threatened by measurement error – ignore this for now

Multiple regression is provably unreliable

for causal inference unless:

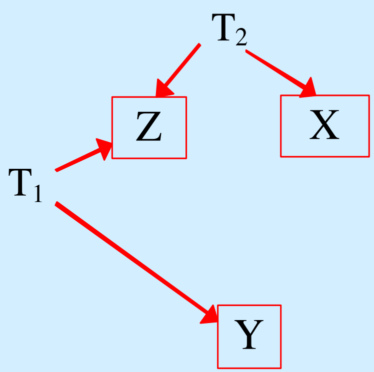
- X prior to Y
- X, **Z**, and Y are causally sufficient (no confounding)

Truth	Regression Y: outcome X, Z, Explanatory	Alternative?



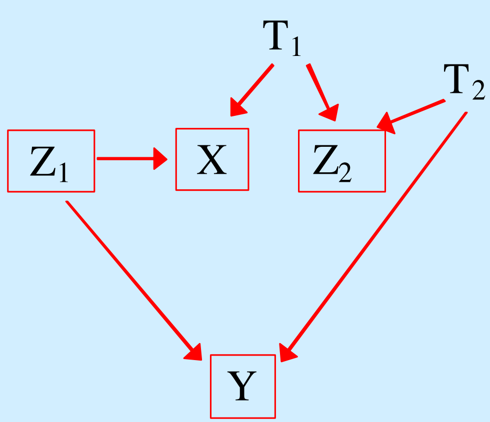
$\beta_X = 0$

$\beta_Z \neq 0$



$\beta_X \neq 0$

$\beta_Z \neq 0$



$\beta_X \neq 0$

$\beta_{Z1} \neq 0$

$\beta_{Z2} \neq 0$

Better Methods Exist

Causal Model Search (since 1988):

- Provably Reliable
- Provably Rumsfeld

Tetrad Demo