SOC 103M

Multiple causation/
Multivariate analysis

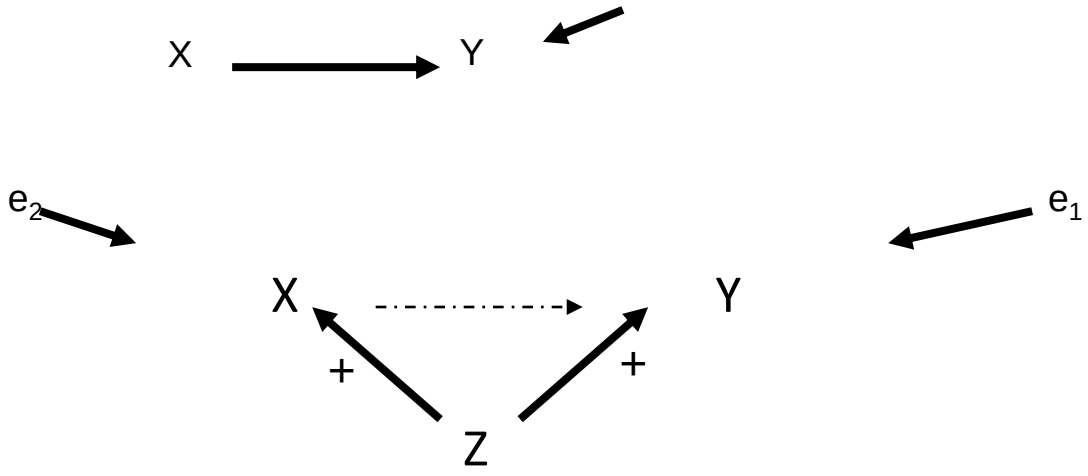
John Stuart Mill's 3 Main Criteria of Causation (recall)

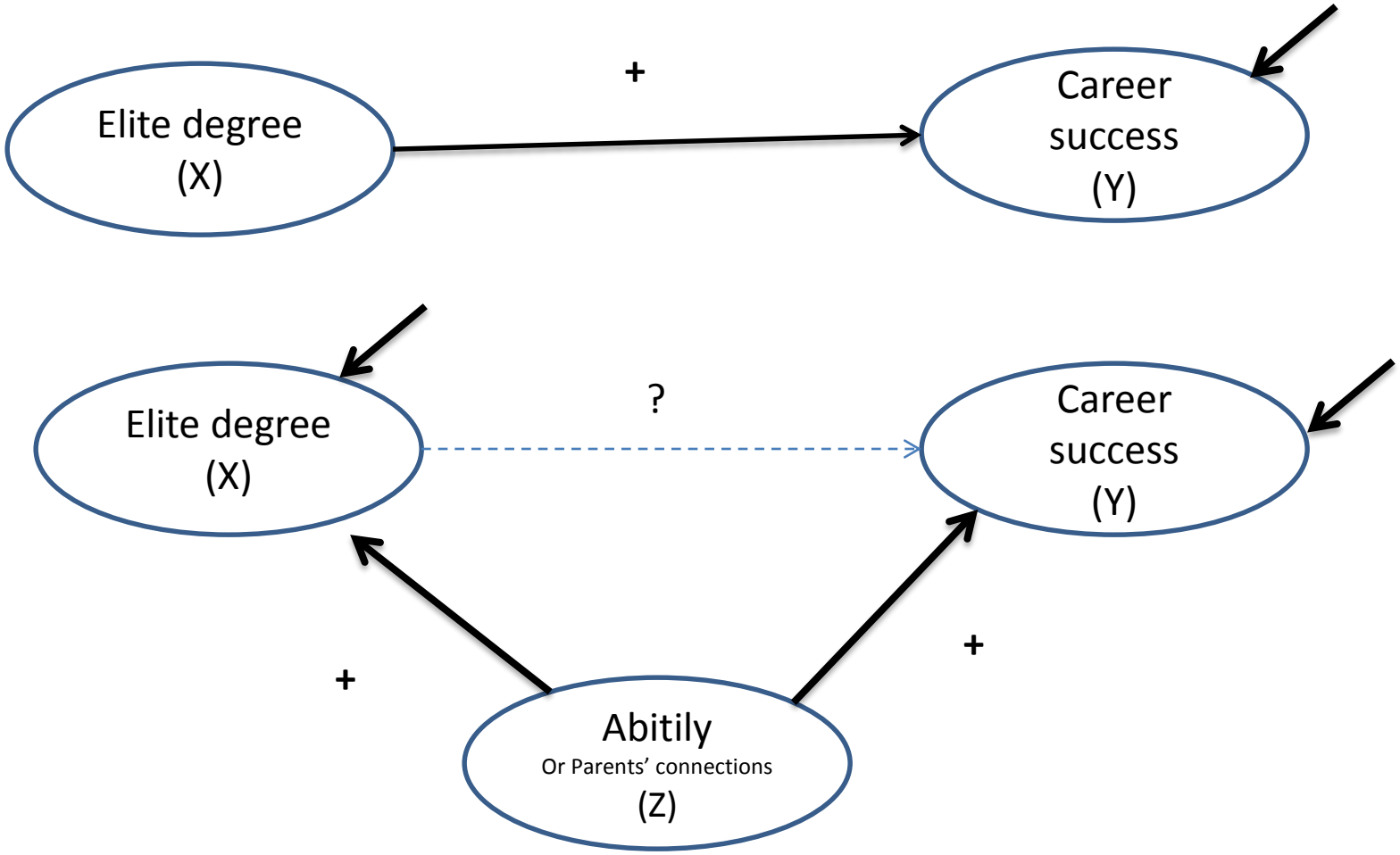
- ***#1. Empirical Association***
- ***#2. Appropriate Time Order***
- ***#3. Non-Spuriousness (Excluding other Forms of Causation)***
  - Mill tells us that even individual causal relationships cannot be established without multivariate analysis (#3).

- Suppose we suspect X causes Y

- Suppose we establish that X is related to Y (#1) and X precedes Y (#2).

- But what if both X and Y are the result of Z a third variable:
  - Both Success (Y) and an Elite Degree (X) are driven by Abitliy or Parents' connections (Z)




Elite degree
(X)
+
Career
success
(Y)
Elite degree
(X)
?
Career
success
(Y)
+
+
Abitily
Or Parents' connections
(Z)

Excluding other Forms of Causation or Eliminating Confounding Factors

- How to establish the independent effect of a variable and exclude/control for a third confounding variable (Z)?
  - Physical control
    - Glass beaker, temperature control etc.
  - Randomization
    - Making treatment and control groups identical in the aggregate
      - Controls for ALL variables (all Zs)
  - Statistical control
    - Separating observations that are identical with respect to some variable Z
      - Controls for ONE variable at a time
    - Logic: if the cases within the group are the same with respect to the control variable, differences within the group cannot be due to differences in the control variable (because there are no such differences)

True Experiments

- Two comparison groups
- Variation in the independent variable (X) before assessment of change in the dependent variable (Y)
- Random assignment to the two (or more) comparison groups
  - Plus
- Identification of a causal mechanism (no black box)
- Control over the context of the experiment
-
- ***A) Experimental (Treatment) and Comparison and Control Groups***
-
- Experimental Group: receives the treatment (X)
- Comparison Group: receives the treatment but in a different quantitiy
- Control Group: does not receive the treatment at all

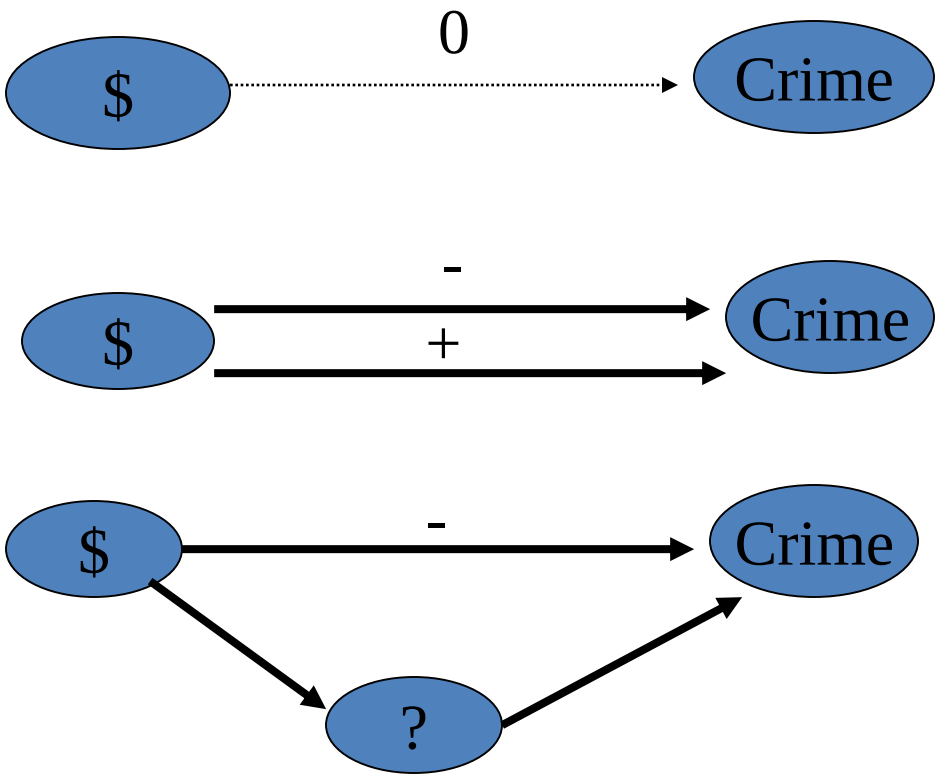
Example: Rossi et al.'s TARP (Transitional Aid Research Project) experiment in Texas and Georgia

- Why do people commit crime?
- Why experiments:
- ecological fallacy
- small variance of exogeneous variables
- multi-collinearity
- bi-directional causation
-
- Replication (Georgia)
-
- They present their theoretical model: money will lower recidivism because:
- declining marginal utility of income
- opportunity cost

Example: Rossi et al.'s TARP (Transitional Aid Research Project) experiment in Texas and Georgia (cont.)

- Six groups (four treatment one control + another not interviewed).
-
- Treatment: payment and job-counseling and placement.
-
- Group 1 26 weeks 100% tax
- Group 2 13 weeks 100% tax
- Group 3 13 weeks 25% tax
- Group 4 job placement only
- Group 5 controls interviewed
- Group 6 not-interviewed controls
-

Suppression



The Importance of Temporal Sequence

The temporal position of Z vis-à-vis X					
	No change/	Zero or statistically not significant	Weaker but statistically significant	Uneven among the categories of Z	Stronger than the unconditional effect
Antecedent variable (Z precedes both X and Y	Z is not a factor	Spuriousness	X is a factor but some of its original effect is spurious	Statistical Interaction	Suppression
Intervening variable (Z precedes Y but not X)	Z is not a factor	Explanation (X has only indirect effect)	X is a factor and it effects Y both indirectly through Z and directly (or through other variables missing from the model)	Statistical Interaction	Suppression

Does Education Make You Happy? The Bivariate Relationship

GENERAL HAPPINESS * Degree recoded Crosstabulation

			Degree recoded			
			LT HIGH SCHOOL	HIGH SCHOOL	AT LEAST SOME COLLEGE	Total
GENERAL HAPPINESS	VERY HAPPY	Count	54	249	148	451
		% within Degree recoded	25.4%	30.2%	32.9%	30.3%
	PRETTY HAPPY	Count	121	482	258	861
		% within Degree recoded	56.8%	58.4%	57.3%	57.9%
	NOT TOO HAPPY	Count	38	94	44	176
		% within Degree recoded	17.8%	11.4%	9.8%	11.8%
Total		Count	213	825	450	1488
		% within Degree recoded	100.0%	100.0%	100.0%	100.0%

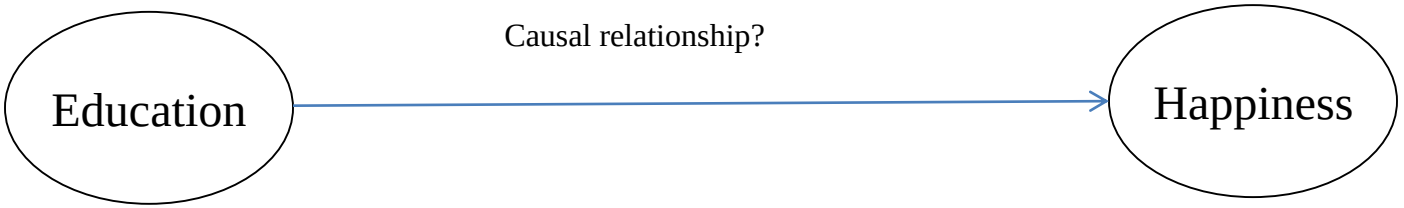
Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Nominal by Nominal	Phi	.086			.026
	Cramer's V	.061			.026
Ordinal by Ordinal	Gamma	-.111	.042	-2.641	.008
N of Valid Cases		1488			

a. Not assuming the null hypothesis.

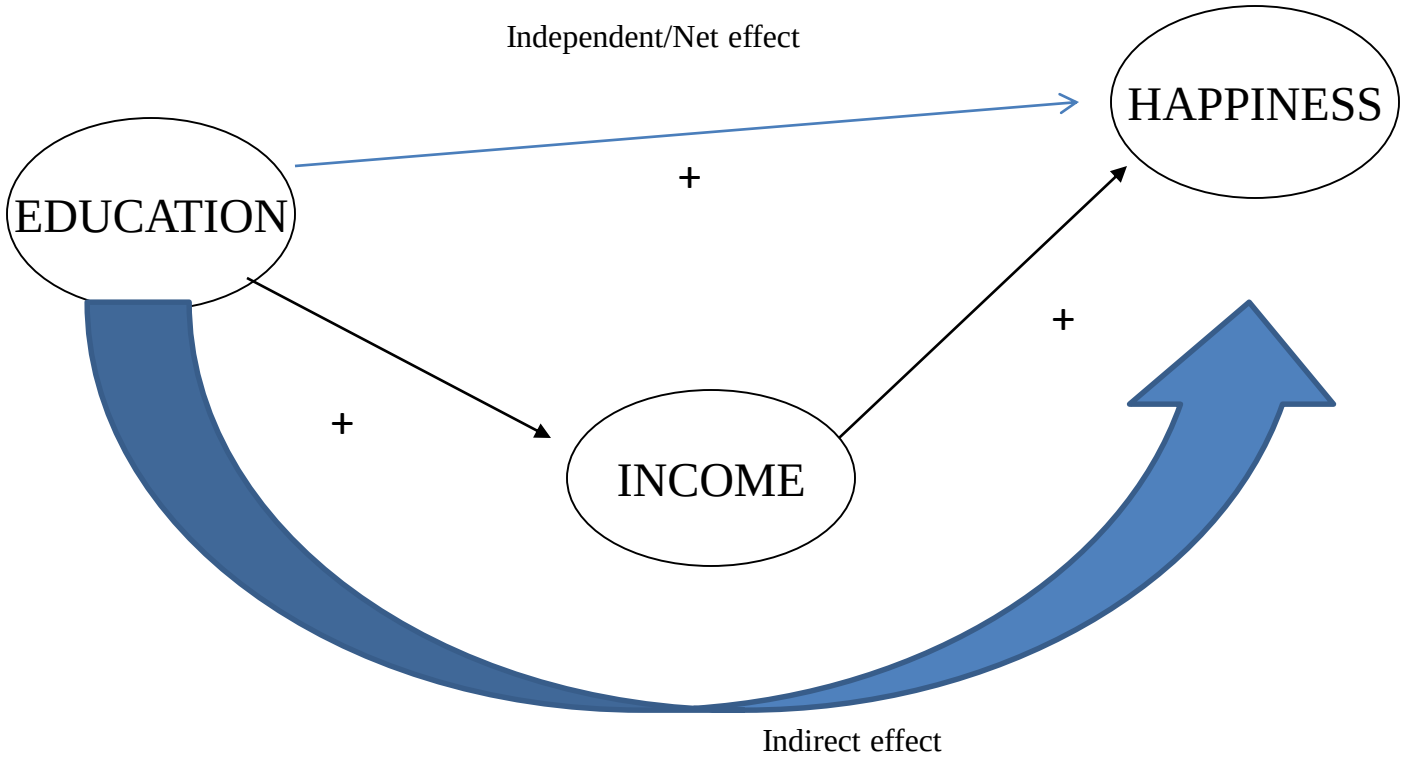
b. Using the asymptotic standard error assuming the null hypothesis.

Education and Happiness



Positive relationship:
More education more happiness

Income as Intervening Variable



GENERAL HAPPINESS * Degree recoded * TOTAL FAMILY INCOME Crosstabulation

TOTAL FAMILY INCOME				Degree recoded			Total
				LT HIGH SCHOOL	HIGH SCHOOL	AT LEAST SOME COLLEGE	
LOWER	GENERAL HAPPINESS	VERY HAPPY	Count	27	66	15	108
			% within Degree recoded	22.7%	23.7%	22.4%	23.2%
		PRETTY HAPPY	Count	66	165	41	272
			% within Degree recoded	55.5%	59.1%	61.2%	58.5%
		NOT TOO HAPPY	Count	26	48	11	85
			% within Degree recoded	21.8%	17.2%	16.4%	18.3%
	Total		Count	119	279	67	465
			% within Degree recoded	100.0%	100.0%	100.0%	100.0%
MIDDLE	GENERAL HAPPINESS	VERY HAPPY	Count	14	81	40	135
			% within Degree recoded	37.8%	30.8%	30.1%	31.2%
		PRETTY HAPPY	Count	19	157	79	255
			% within Degree recoded	51.4%	59.7%	59.4%	58.9%
		NOT TOO HAPPY	Count	4	25	14	43
			% within Degree recoded	10.8%	9.5%	10.5%	9.9%
	Total		Count	37	263	133	433
			% within Degree recoded	100.0%	100.0%	100.0%	100.0%
HIGHER	GENERAL HAPPINESS	VERY HAPPY	Count	1	67	76	144
			% within Degree recoded	6.7%	39.0%	35.8%	36.1%
		PRETTY HAPPY	Count	12	93	123	228
			% within Degree recoded	80.0%	54.1%	58.0%	57.1%
		NOT TOO HAPPY	Count	2	12	13	27
			% within Degree recoded	13.3%	7.0%	6.1%	6.8%
	Total		Count	15	172	212	399
			% within Degree recoded	100.0%	100.0%	100.0%	100.0%

The Relationship Between Education and Happiness
Controlling for Income

Symmetric Measures

TOTAL FAMILY INCOME			Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
LOWER	Nominal by Nominal	Phi	.056			.830
		Cramer's V	.040			.830
	Ordinal by Ordinal	Gamma	-.049	.076	-.643	.520
	N of Valid Cases		465			
MIDDLE	Nominal by Nominal	Phi	.051			.893
		Cramer's V	.036			.893
	Ordinal by Ordinal	Gamma	.049	.085	.579	.563
	N of Valid Cases		433			
HIGHER	Nominal by Nominal	Phi	.130			.149
		Cramer's V	.092			.149
	Ordinal by Ordinal	Gamma	-.045	.089	-.508	.611
	N of Valid Cases		399			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

The Importance of Temporal Sequence

The temporal position of Z vis-à-vis X					
	No change/	Zero or statistically not significant	Weaker but statistically significant	Uneven among the categories of Z	Stronger than the unconditional effect
Antecedent variable (Z precedes both X and Y	Z is not a factor	Spuriousness	X is a factor but some of its original effect is spurious	Statistical Interaction	Suppression
Intervening variable (Z precedes Y but not X)	Z is not a factor	Explanation (X has only indirect effect)	X is a factor and it effects Y both indirectly through Z and directly (or through other variables missing from the model)	Statistical Interaction	Suppression