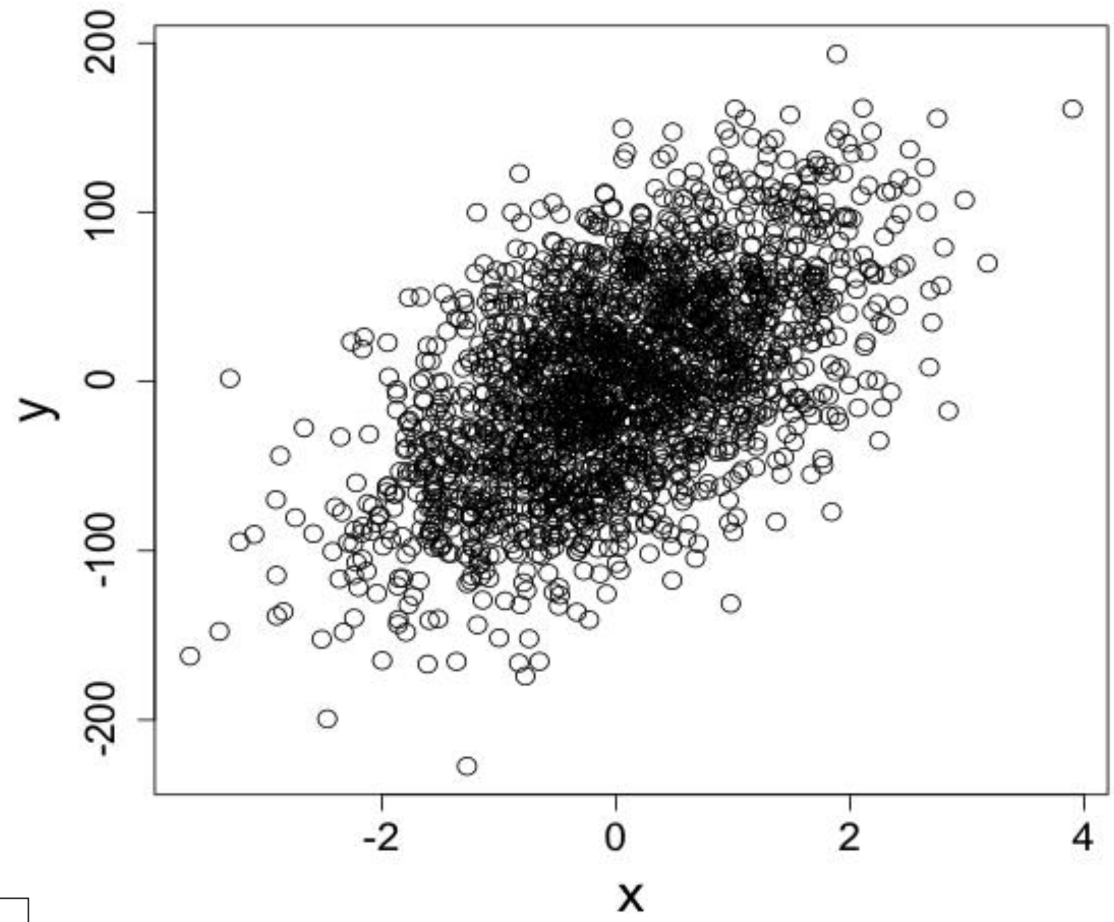
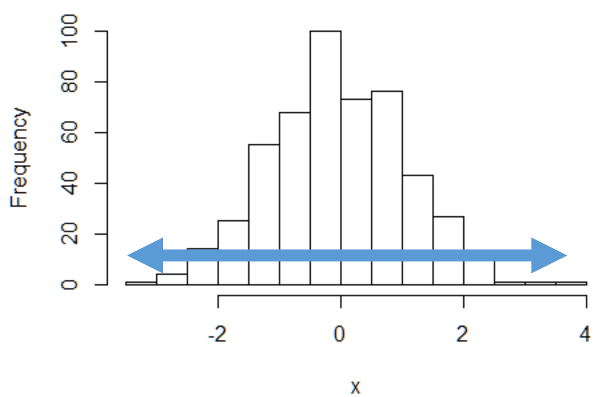
5. Path Coefficients

5. Coefficients. Covariance and correlation

$$VAR_y = \frac{\sum (y_i - \bar{y})^2}{n-1}$$

$$SD_y = \sqrt{VAR_y}$$



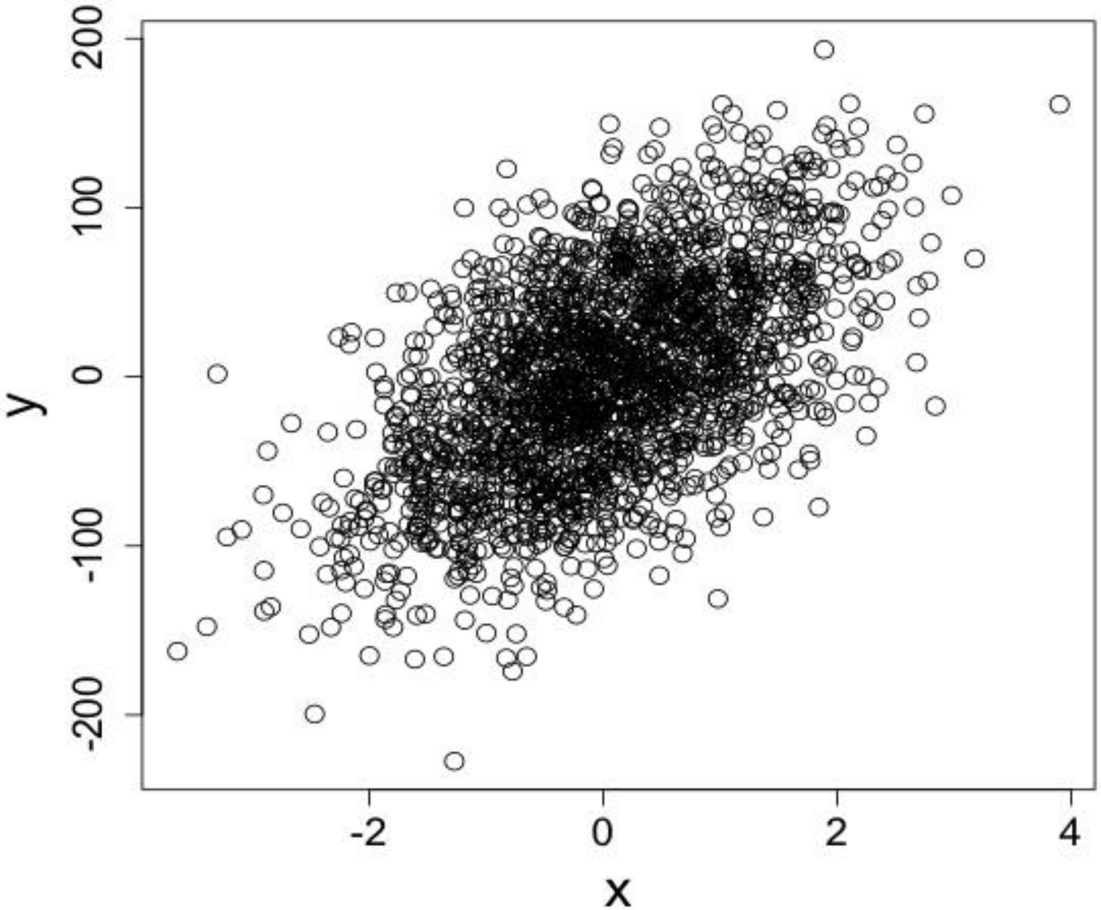


$$VAR_x = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

5. Coefficients. Covariance and correlation

$$COV_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$r_{xy} = \frac{COV_{xy}}{(SD_x \times SD_y)}$$



Covariances *are* correlations when variables are standardized

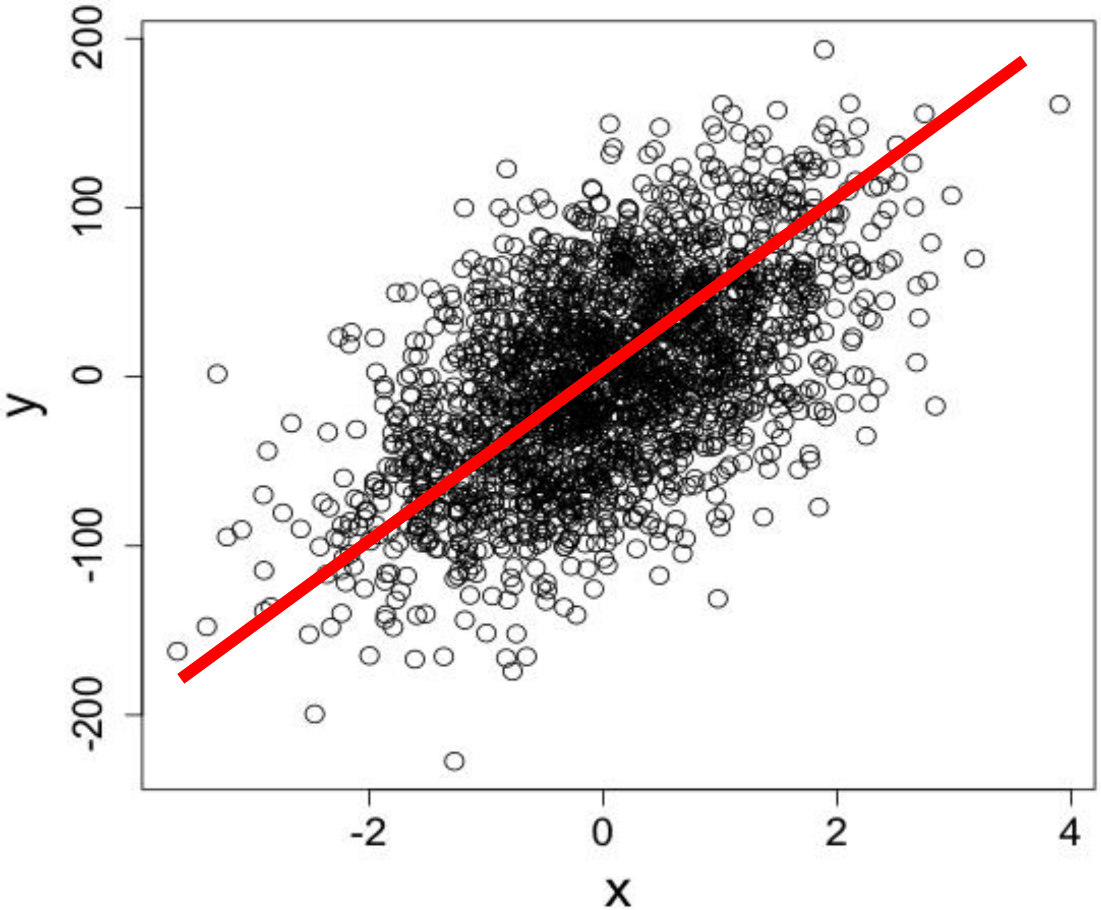
(Z-transformed: subtract the mean and divide by the SD)

5. Coefficients. Covariance and correlation

"rise"

$$\beta_{xy} = \frac{COV_{xy}}{VAR_x}$$

"run"



$$b_{xy} = \frac{COV_{xy}}{VAR_x} = \frac{r_{xy}}{1} = r_{xy}$$

When variables are standardized, the coefficient *is* the correlation for simple regressions

5. Coefficients. Standardization

- *Unstandardized coefficient* = absolute strength of the pathway
  - “ An 1 unit change in *X* results in some unit change in *Y* ”

- *Standardized coefficient* = relative strength of the pathway
  - “ A 1 standard deviation change in *X* results in some standard deviation change in *Y* ”

$$b = B_{xy} * \frac{sd(x)}{sd(y)}$$

5. Coefficients. Standardization

Unstandardized	Standardized
Good for prediction: coefficients are in raw units	Good for ranking: coefficients are in equivalent units
Has direct real world meaning	Less clear real world meaning
Can be compared across pathways or models that have identical units	Can be compared across all pathways in the same model or same population; CANNOT be compared across different statistical populations (different population-level variances)

5. Coefficients. Covariance and correlation

We often use covariances to fit models, but standardized covariances – i.e. correlations – for interpretation.

Raw Covariance Matrix

	x_1	x_2	y_1
x_1	0.81		
x_2	0.87	1.63	
y_1	0.88	1.80	4.98

variance

covariance

Standardized Covariance Matrix

	x_1	x_2	y_1
x_1	1.0		
x_2	0.76	1.0	
y_1	0.44	0.63	1.0

correlation

5. Coefficients. Standardization

```
# Generate random data
set.seed(1)

data <- data.frame(x1 = rnorm(100))

data$x2 <- data$x1 + runif(100, 0, 3)

data$y1 <- data$x2 + runif(100, 0, 6)

data$y2 <- data$x2 + runif(100, 0, 9)

# Standardized coefficients: Bxy * sd(x) / sd(y)
mod <- lm(y2 ~ y1, data)

# `coefs` returns the coefficient table (both standardized and
unstandardized)

piecewiseSEM::coefs(mod)


  Response Predictor Estimate Std.Error DF Crit.Value P.Value Std.Estimate
1       y2        y1   0.3433    0.1294 98     2.6528  0.0093       0.2588 **
```

5. Coefficients. Standardization

```
# `unstdCoefs` and `stdCoefs` return unrounded coefficients
stdCoefs(mod)

 Response Predictor  Estimate Std.Error DF Crit.Value     P.Value Std.Estimate
1       y2        y1 0.3432863 0.1294042 98   2.652821 0.009312599    0.2588427

BetaStd <- stdCoefs(mod)$Estimate * sd(data$y1) / sd(data$y2)

# Compare manually standardized to automatically standardized output
BetaStd; stdCoefs(mod)$Std.Estimate

[1] 0.2588427
[1] 0.2588427
```

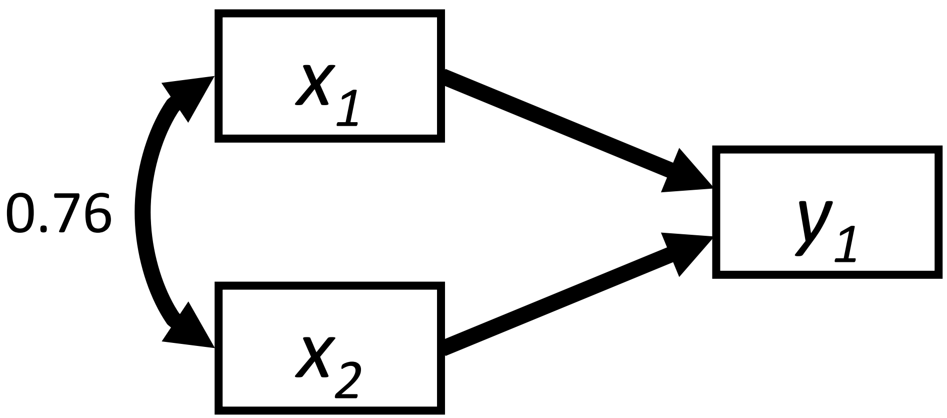
5. Coefficients. Standardization

```
# The same as scaling the data beforehand and retrieving the raw
coefficients
data.scaled <- as.data.frame(apply(data, 2, scale))

mod2 <- lm(y2 ~ y1, data.scaled)

stdCoefs(mod2) # Estimate and Std.Estimate are the same

  Response Predictor  Estimate Std.Error DF Crit.Value     P.Value Std.Estimate
1       y2        y1 0.2588427 0.0975726 98   2.652821 0.009312599    0.2588427
```

The 8 Rules of Path Coefficients

5. Coefficients. Rule #1 of path coefficients

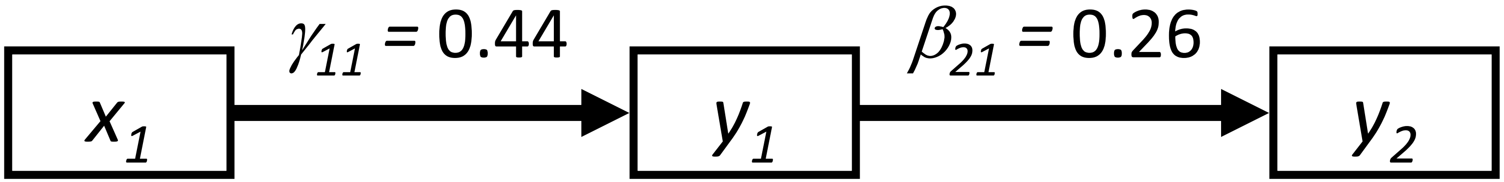
First Rule of Path Coefficients: path coefficients for unanalyzed relationships (double-headed arrows) between exogenous variables are simply the correlations (standardized form) or covariances (unstandardized form)



	x_1	x_2	y_1
x_1	1.0		
x_2	0.76	1.0	
y_1	0.44	0.63	1.0

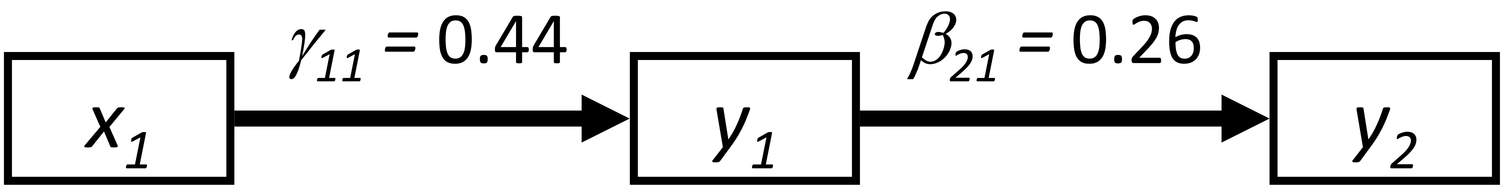
5. Coefficients. Rule #2 of path coefficients

Second Rule of Path Coefficients: when variables are connected by a *single* causal path, the (standardized) path coefficient is simply the correlation coefficient



	x_1	y_1	y_2
x_1	1.0		
y_1	0.44	1.0	
y_2	0.31	0.26	1.0

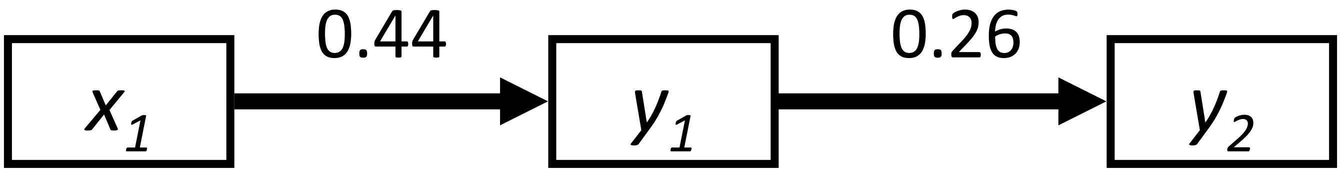
5. Coefficients. Rule #2 of path coefficients




```
cor(data[, -2])

# Path 1
mody1.x1 <- lm(y1 ~ x1, data)

stdCoefs(mody1.x1)$Std.Estimate; cor(data[, c("y1", "x1")])[2, 1]

[1] 0.4400536
[1] 0.4400536

# Path 2
mody2.y1 <- lm(y2 ~ y1, data)

stdCoefs(mody2.y1)$Std.Estimate; cor(data[, c("y2", "y1")])[2, 1]

[1] 0.2588427
[1] 0.2588427
```

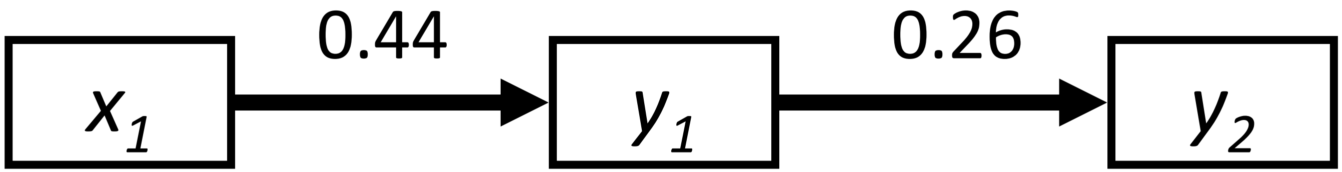
5. Coefficients. Rule #3 of path coefficients

Third Rule of Path Coefficients: strength of a compound path is the product of the (standardized) coefficients along the path.



If the indirect path from x_1 to y_2 <u>equals the correlation</u> between x_1 and y_2, we say x_1 and y_2 are *conditionally independent*.

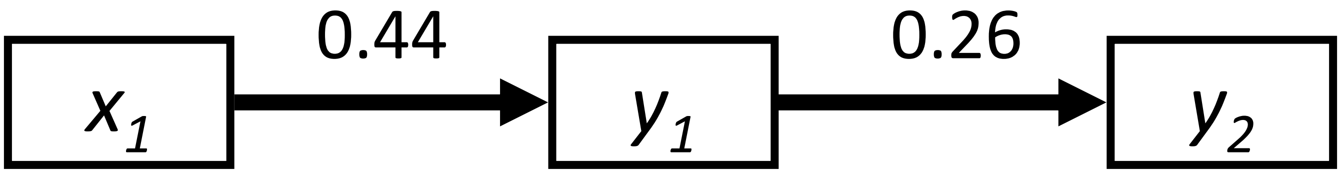
5. Coefficients. Rule #3 of path coefficients




```
stdCoefs(mody1.x1)$Std.Estimate * stdCoefs(mody2.y1)$Std.Estimate;
cor(data[, c("y2", "x1")])[2, 1]

[1] 0.1139046
[1] 0.3138744

# Wait a minute...!
```

5. Coefficients. Rule #3 of path coefficients

What does it mean when two separated variables are *not* conditionally independent?

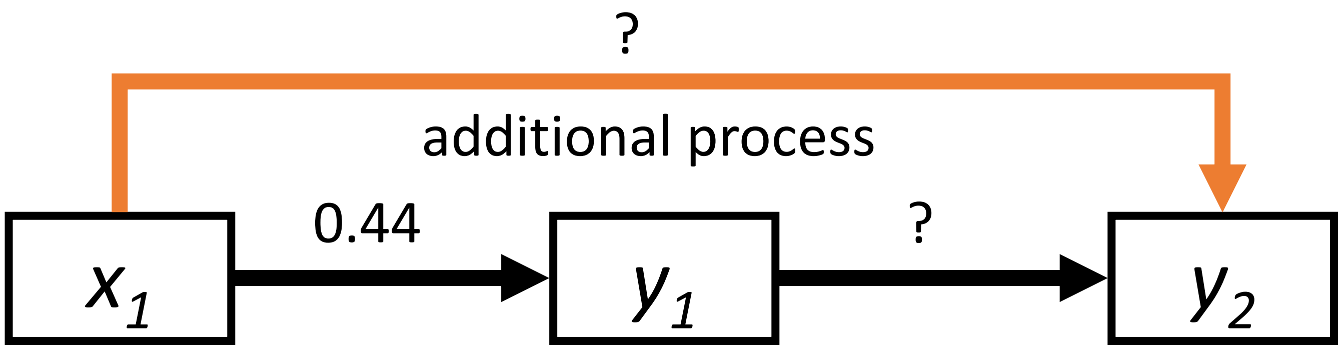


0.44 * 0.26 = 0.11, which is <u>not equal</u> to 0.31

	x_1	y_1	y_2
x_1	1.0		
y_1	0.44	1.0	
y_2	0.31	0.26	1.0

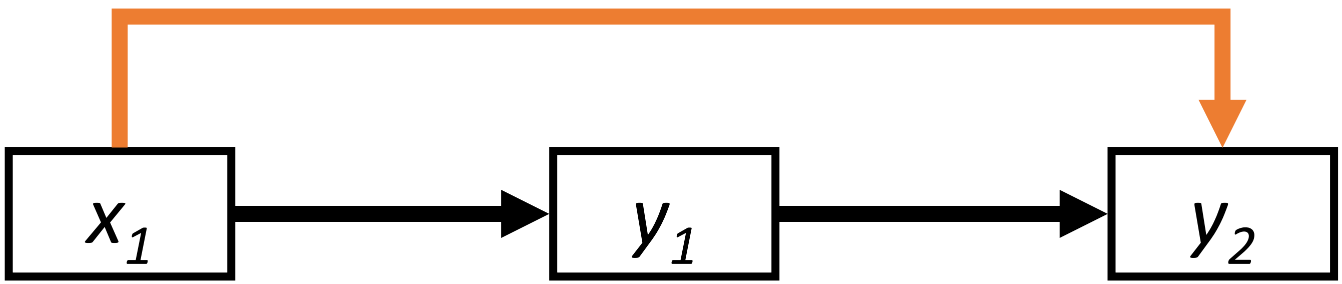
5. Coefficients. Rule #4 of path coefficients

The inequality implies that the true model is:



Fourth Rule of Path Coefficients: when variables are connected by more than one causal pathway, the path coefficients are "partial" regression coefficients.

5. Coefficients. What is a partial coefficient?



Direct correlation

Indirect correlations

$$\gamma_{21} = \frac{r_{x_1 y_2} - (r_{x_1 y_1} \times r_{y_1 y_2})}{1 - r^2_{x_1 y_1}}$$

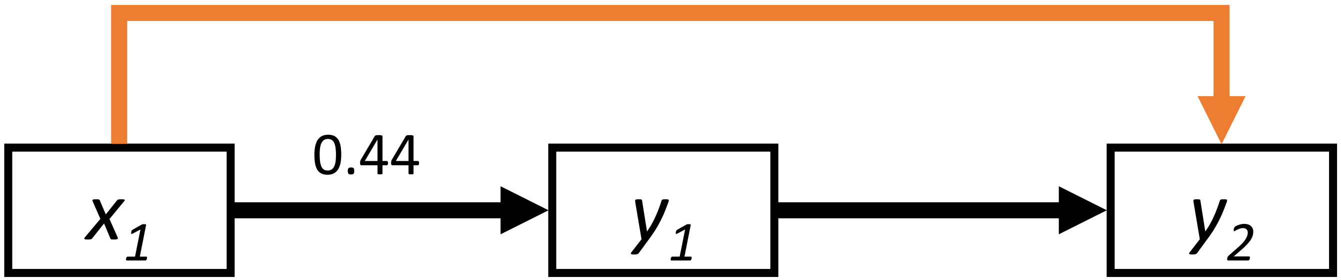
Shared variance between predictors

5. Coefficients. What is a partial coefficient?

$$\gamma_{21} = \frac{r_{x_1 y_2} - (r_{x_1 y_1} \times r_{y_1 y_2})}{1 - r^2_{x_1 y_1}}$$

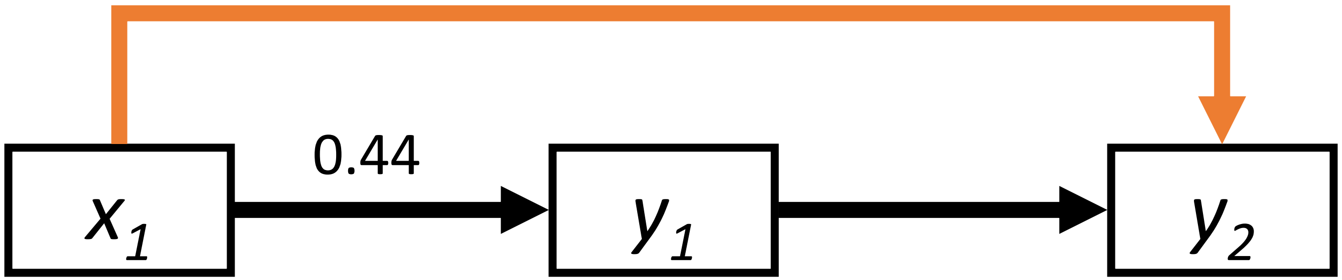
	x_1	y_1	y_2
x_1	1.0		
y_1	0.44	1.0	
y_2	0.31	0.26	1.0

$$\beta_{12} = \frac{(0.31 - (0.26 * 0.44))}{1 - 0.44^2} = 0.25$$



5. Coefficients. What is a partial coefficient?

$$\beta_{12} = \frac{(0.31 - (0.26 * 0.44))}{1 - 0.44^2} = 0.25$$




```
# Path 1
mody2.x1 <- lm(y2 ~ y1 + x1, data)

Gamma.y2.x1 <- (cor(data$y2, data$x1) - (cor(data$y2, data$y1) *
cor(data$y1, data$x1))) /
  (1 - cor(data$y1, data$x1) ^ 2)

stdCoefs(mody2.x1)[2, 8]; Gamma.y2.x1

[1] 0.2479929
[1] 0.2479929
```

5. Coefficients. What is a partial coefficient?

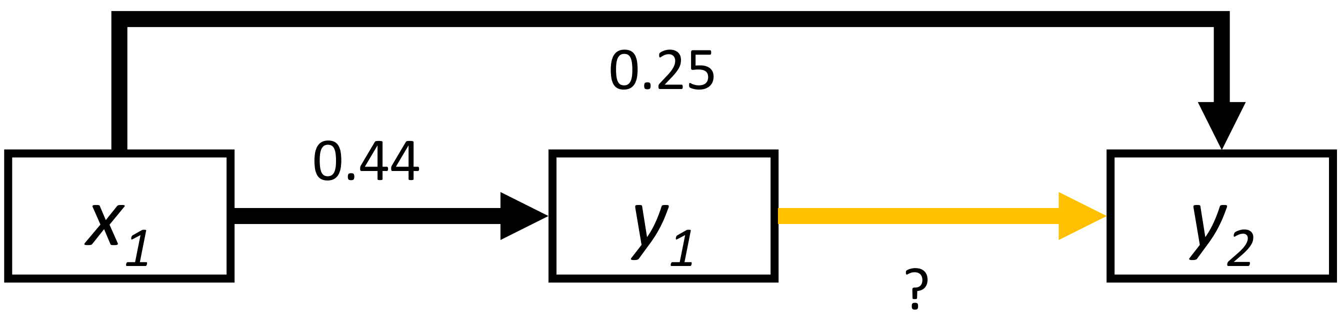
Direct correlation

Indirect correlations

$$\beta_{21} = \frac{r_{y_1 y_2} - (r_{x_1 y_1} \times r_{x_1 y_2})}{1 - r^2_{x_1 y_1}}$$

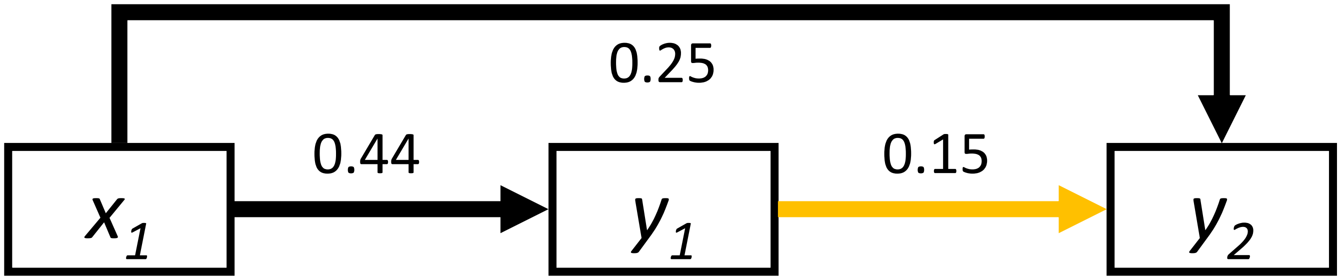
Shared variance between predictors

	x_1	y_1	y_2
x_1	1.0		
y_1	0.44	1.0	
y_2	0.31	0.26	1.0



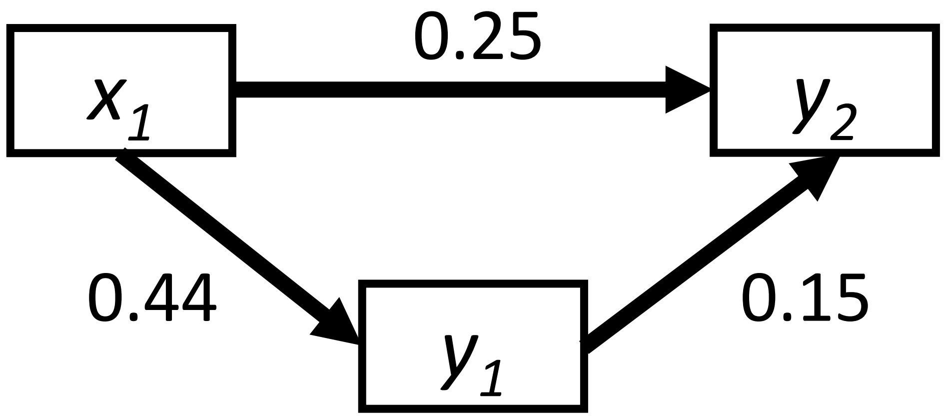
$$\gamma_{21} = \frac{(0.26 - (0.44 * 0.31))}{1 - 0.44^2} = 0.15$$

5. Coefficients. What is a partial coefficient?




```
# Path 2
Gamma.y2.y1 <- (cor(data$y2, data$y1) - (cor(data$y2, data$x1) *
cor(data$y1, data$x1))) /
  (1 - cor(data$y1, data$x1) ^ 2)

stdCoefs(mody2.x1)[1, 8]; Gamma.y2.y1

[1] 0.1497125
[1] 0.1497125
```

5. Coefficients. Statistical control

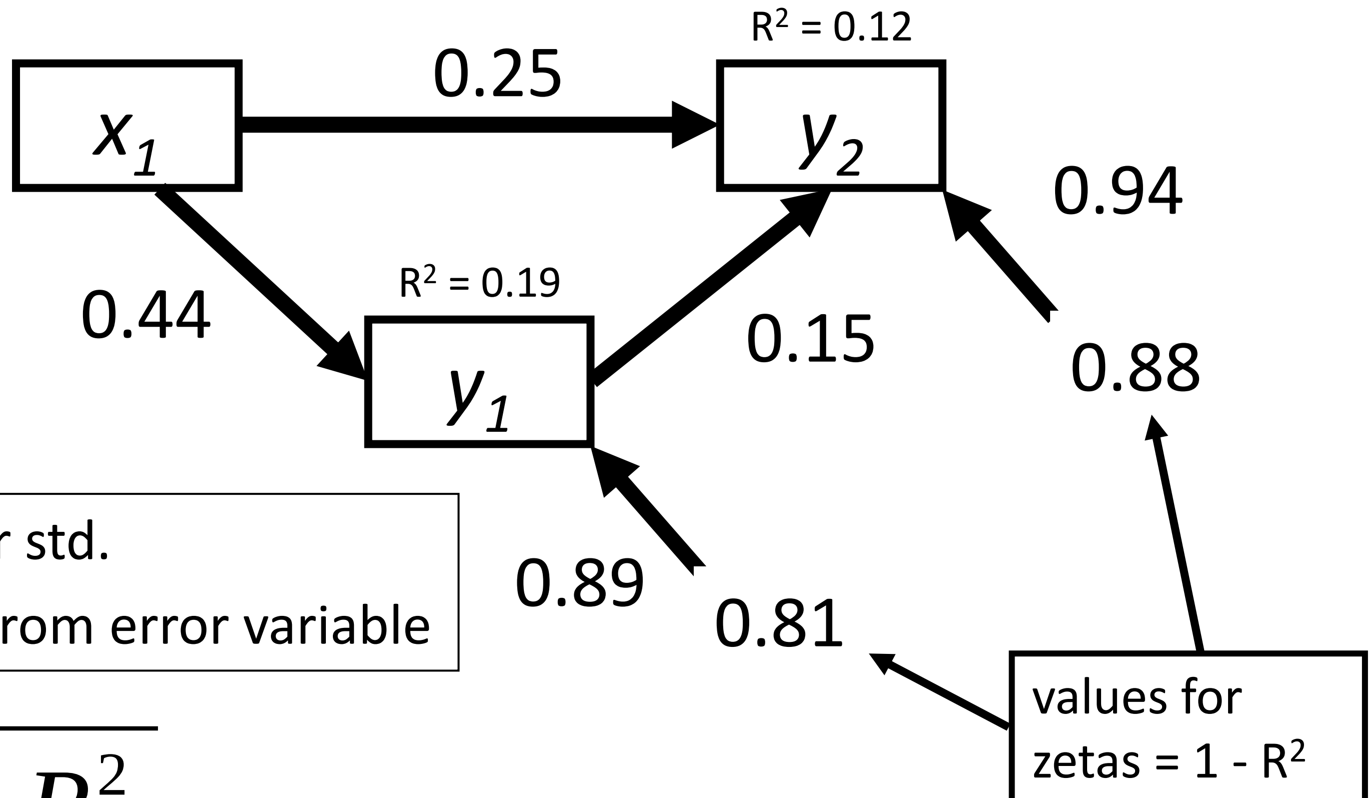


The effect of y_1 on y_2 is controlled for the joint effects of x_1.

With all other variables in model held to their means, how much does a response variable change when a predictor is varied?

5. Coefficients. Rule #5 of path coefficients

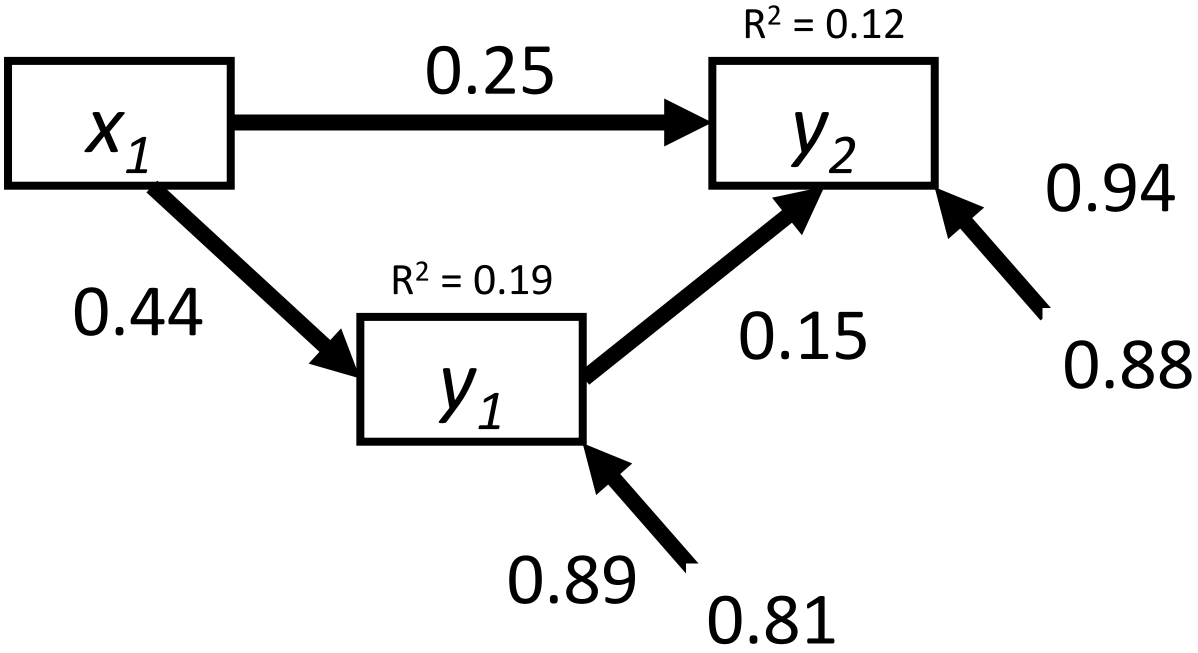
Fifth Rule of Path Coefficients: paths from error variables represent prediction error (influences from other forces).



equation for std.
coefficient from error variable

$$= \sqrt{1 - R^2_{y_i}}$$

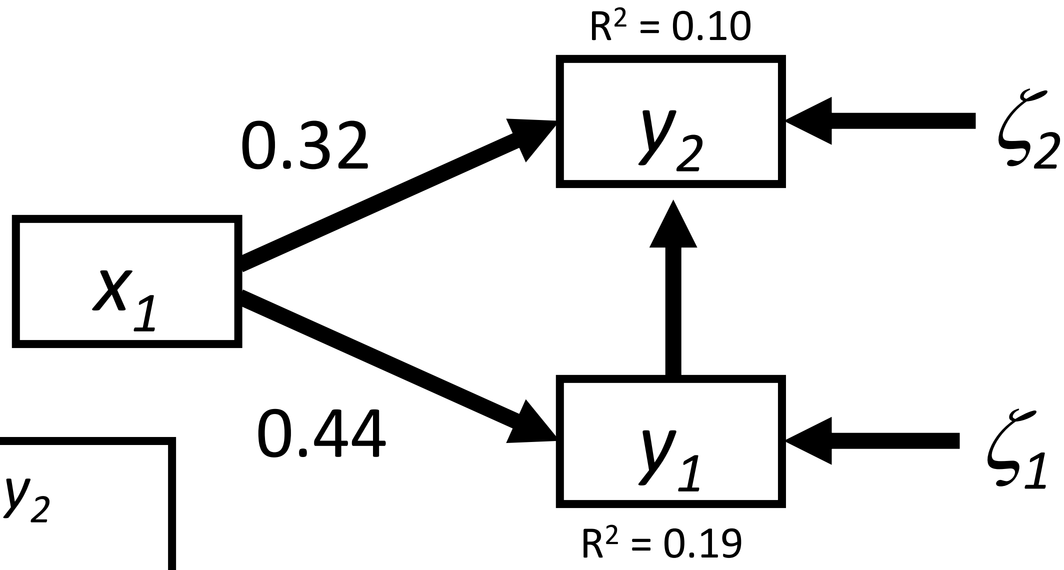
5. Coefficients. Rule #5 of path coefficients




```
Zeta.y1 <- 1 - summary(mody1.x1)$r.squared

sqrt(Zeta.y1)

[1] 0.8979715

Zeta.y2 <- 1 - summary(mody2.x1)$r.squared

sqrt(Zeta.y2)

[1] 0.9398986
```

5. Coefficients. Rule #6 of path coefficients

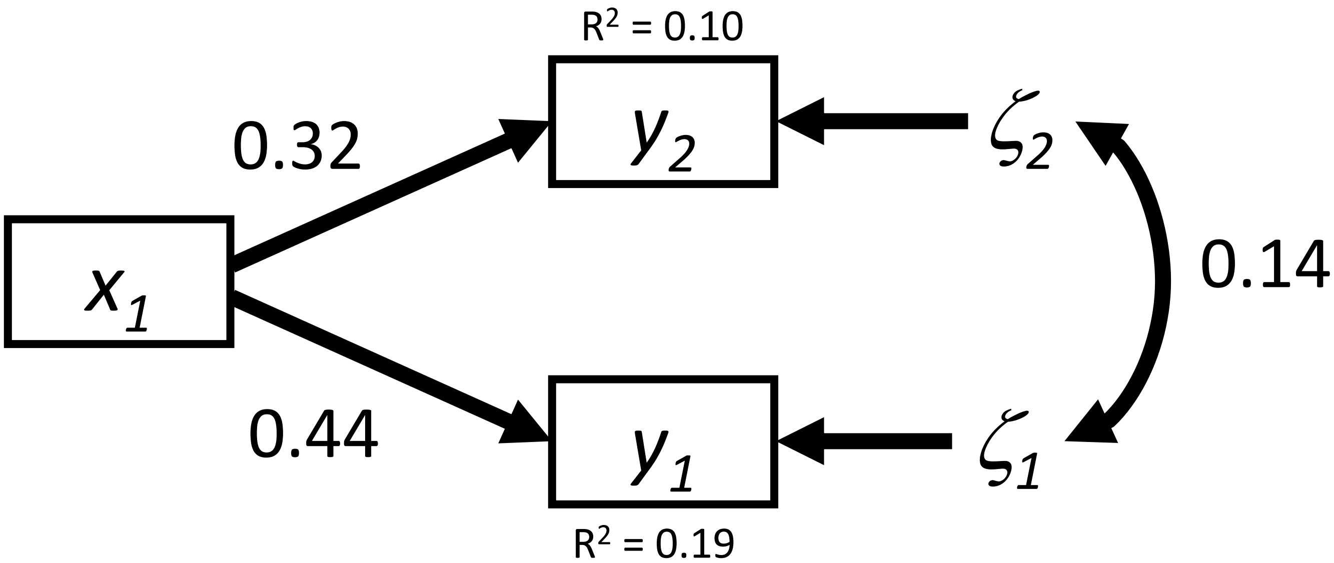
Let's remove the directed path between y_1 and y_2 for moment…



	x_1	y_1	y_2
x_1	1.0		
y_1	0.44	1.0	
y_2	0.31	0.26	1.0

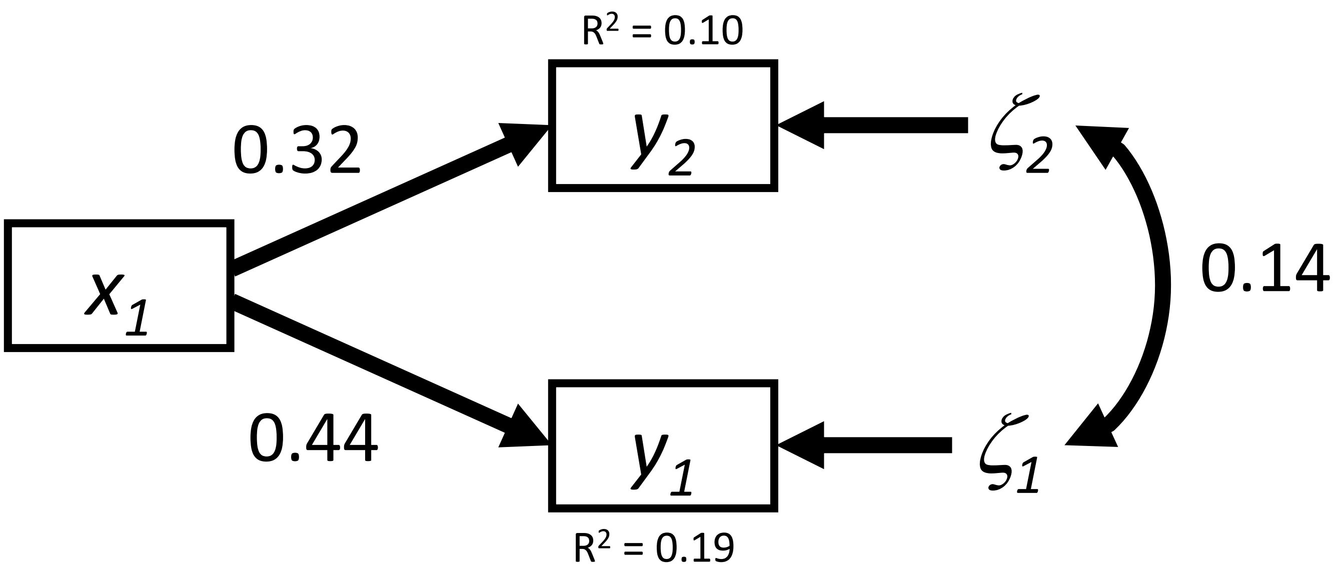
Sixth Rule of Path Coefficients: unanalyzed residual correlations between endogenous variables are partial correlations or covariances.

5. Coefficients. Rule #6 of path coefficients



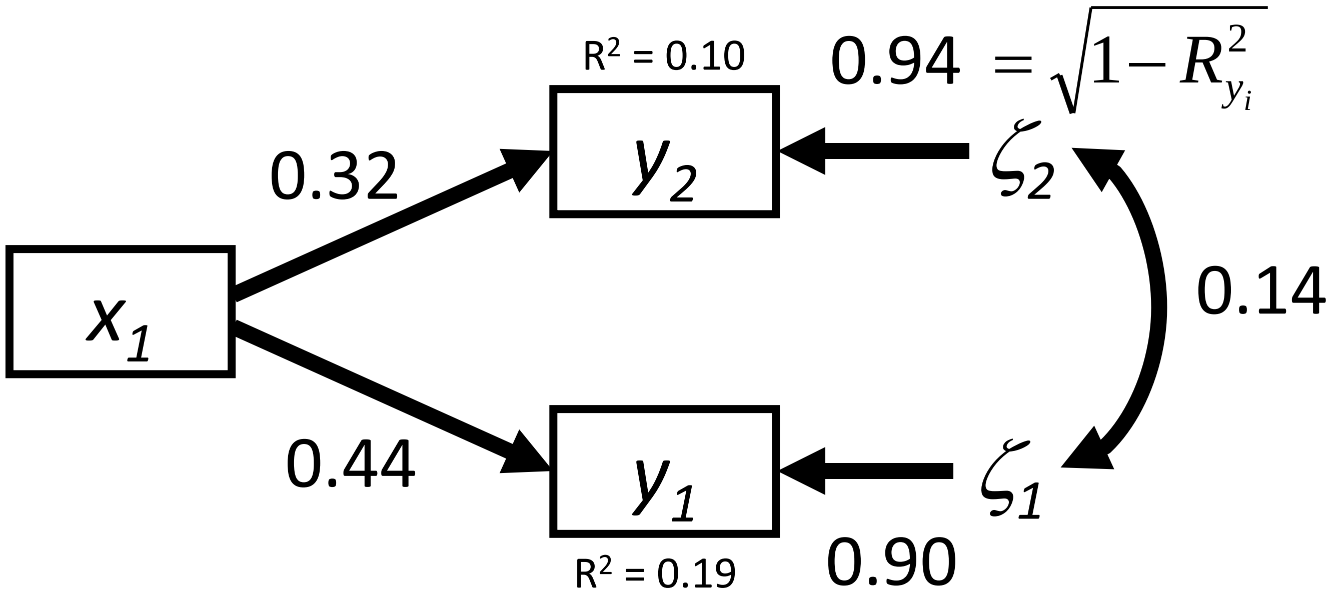

```
(pcor <- cor(
  # effect of x1 on y2
  resid(mody2.x1),
  # effect of x1 on y1
  resid(mody1.x1)
))

[1] 0.1415931
```

5. Coefficients. Rule #6 of path coefficients




```
# Can also use function from piecewiseSEM
partialCorr(y1 %~~% y2, list(mody2.x1, mody1.x1))

  Response Predictor  Estimate Std.Error DF Crit.Value   P.Value
1     ~~y1      ~~y2 0.1415931        NA 97   1.408723 0.9189427
```

5. Coefficients. Rule #6 of path coefficients

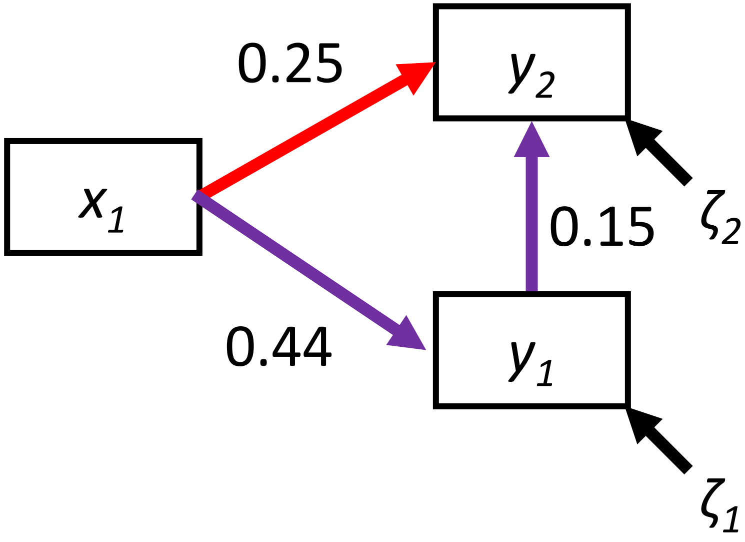


Note that total correlation between y_1 and y_2 = 0.32 * 0.44 + 0.94 * 0.14 * 0.90 = 0.26

	x_1	y_1	y_2
x_1	1.0		
y_1	0.44	1.0	
y_2	0.31	0.26	1.0

5. Coefficients. Rule #7 of path coefficients

Seventh Rule of Path Coefficients: total effect one variable has on another equals the sum of its direct and indirect effects.

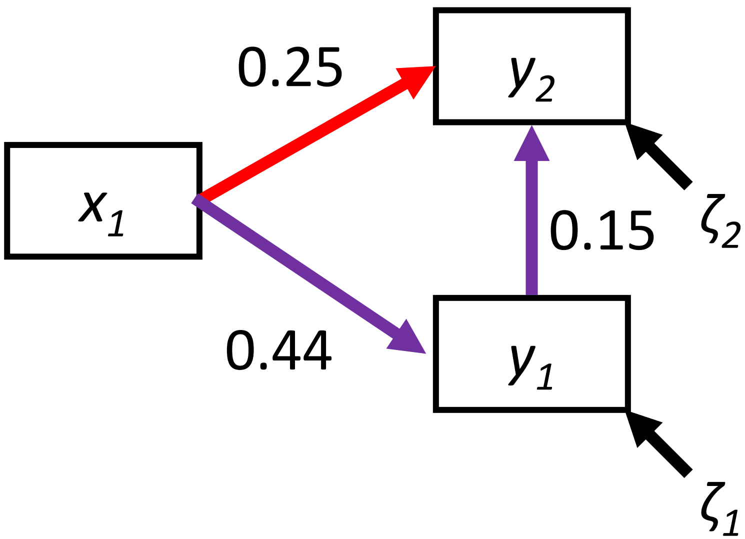


Total Effects:

0.25 + 0.44 * 0.15 = 0.31

5. Coefficients. Rule #8 of path coefficients

Eighth Rule of Path Coefficients: sum of all pathways between two variables (directed and undirected) equals the correlation.

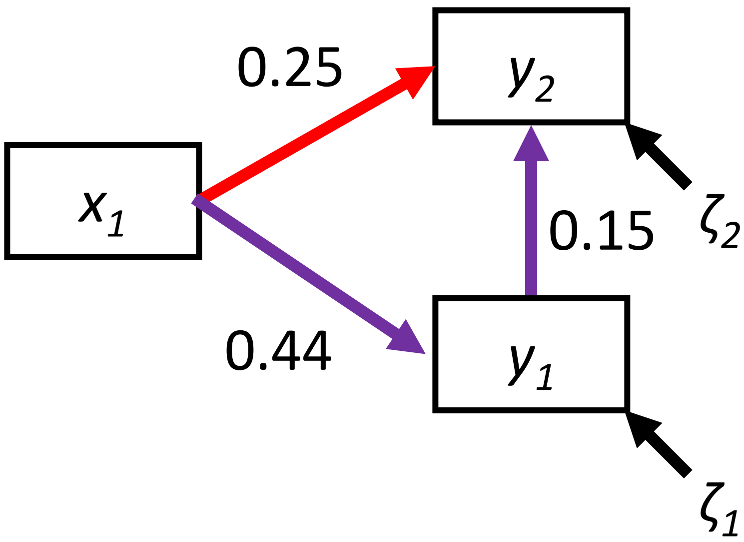


Total Effects:

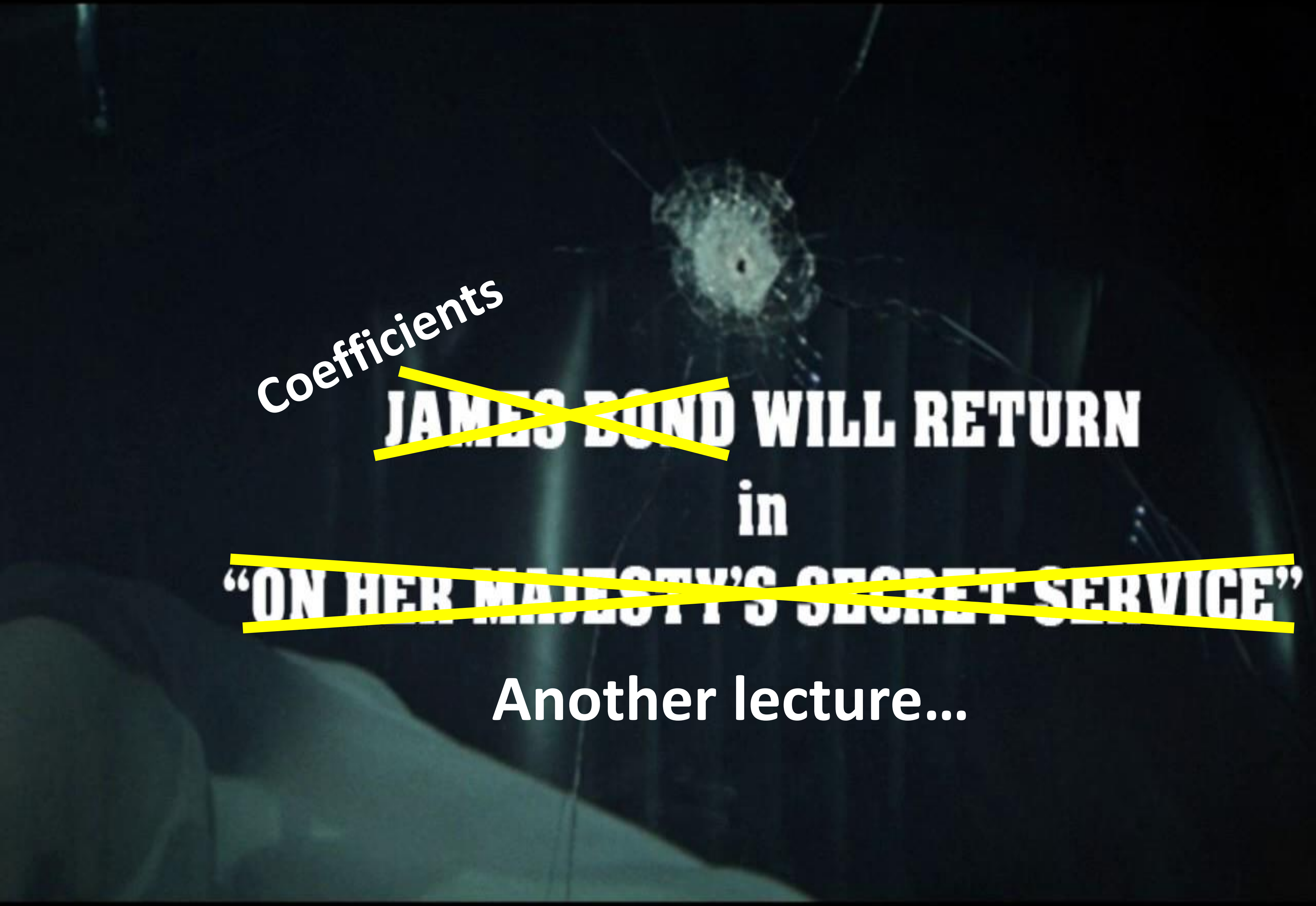
	x_1	y_1	y_2
x_1	1.0		
y_1	0.44	1.0	
y_2	0.31	0.26	1.0

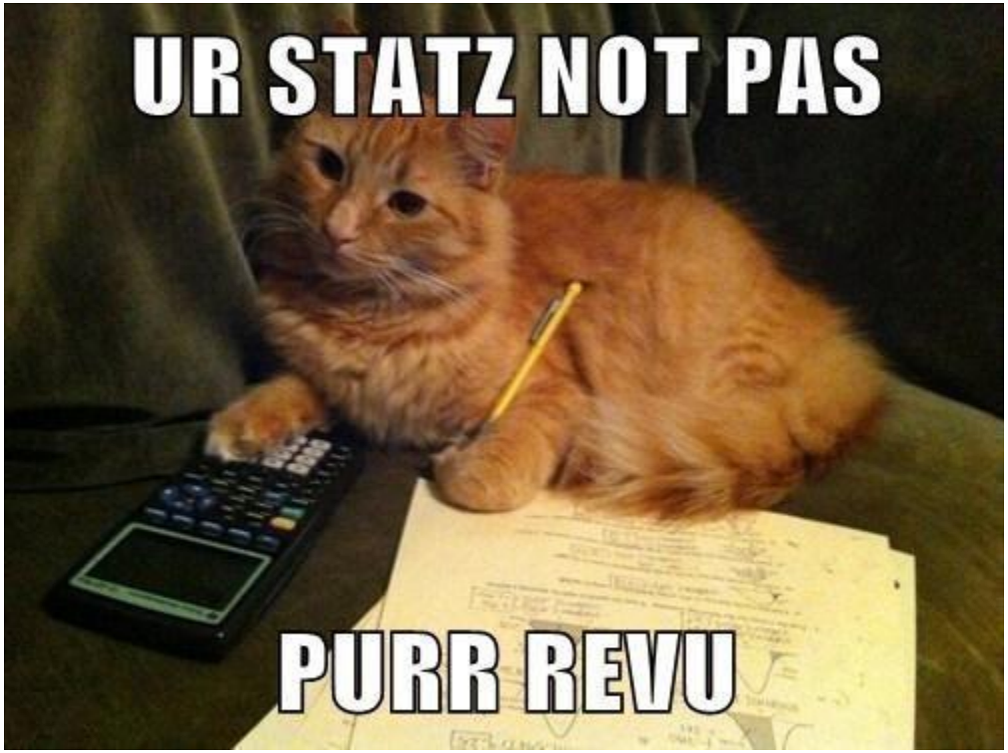
0.25 + 0.44 * 0.15 = 0.31

5. Coefficients. Rule #8 of path coefficients




```
mody2 <- lm(y2 ~ x1 + y1, data)

mody1 <- lm(y1 ~ x1, data)

Gamma.y2.x1 <- stdCoefs(mody2)$Std.Estimate[1] +
(stdCoefs(mody1)$Std.Estimate[1] * stdCoefs(mody2)$Std.Estimate[2])

Gamma.y2.x1; cor(data$x1, data$y2)

[1] 0.3138744
[1] 0.3138744
```


Coefficients
~~JAMES BOND~~ WILL RETURN
in
~~“ON HER MAJESTY’S SECRET SERVICE”~~
Another lecture…


UR STATZ NOT PAS
PURR REVU