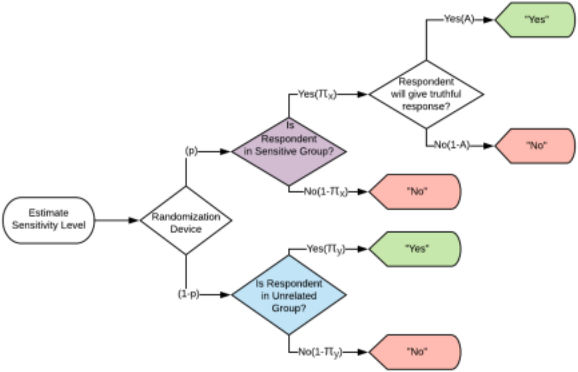
Warner Solution to Untruthful Responding and Mixed Models

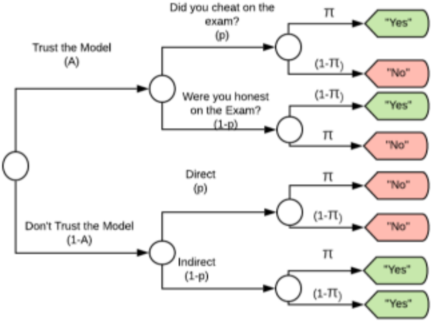
How I Spent My Weekend

Maxwell Lovig



$$\hat{\pi}_g^* = \frac{\hat{P}_y^* - (1-p)\pi_y}{p}$$

Warner's Model



$$\hat{\pi}^*_w = \frac{\hat{P}^*_y - (1-p)\pi_y}{2p-1}$$

Comparison of Variance

Both Models under untruthful responding have the same bias of $\pi(A-1)$

$$Var(\hat{\pi}_g^*) = \frac{\hat{P}_y^*(1-\hat{P}_y^*}{p^2(n-1)}$$

$$Var(\hat{\pi}_w^*) = \frac{\hat{P}_y^*(1-\hat{P}_y^*}{(2p-1)^2(n-1)}$$

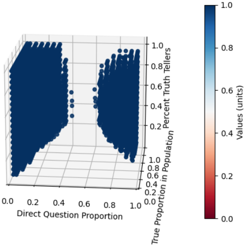
Comparision between simular values

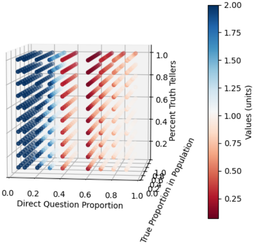
Comparison of Variances Under the Two Models with Untruthful Responding				
p	**A**	$Var(\hat{\pi}_x^*)$	$Var(\hat{\pi}_w^*)$	**Bias** $= \pi(A-1)$
0.2	0.90	0.011873948	0.00128566	-0.03
0.2	0.80	0.011940681	0.001256201	-0.06
0.3	0.90	0.005454442	0.003025251	-0.03
0.3	0.80	0.005481096	0.002995792	-0.06

$$n = 500,\ \pi = 0.3,\ \pi_y = 0.7$$

Though Simulation we can extend this to more numbers!

Turns out Warner is more effective in alot of cases.





$n = 100, \pi_x = .3, \pi_y = .7, A = .8$

Incorporation of Untruthfulness

For Greenberg's Model

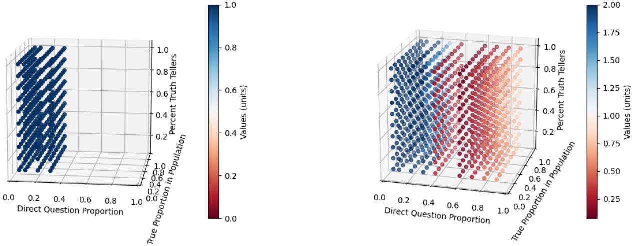
$$\hat{\pi}_g^* = \frac{P(y) - (1-p)\pi_y}{Ap}$$

For Warner's Model

$$\hat{\pi}_w^* = \frac{P(y) - (1-p)}{A(2p-1)}$$

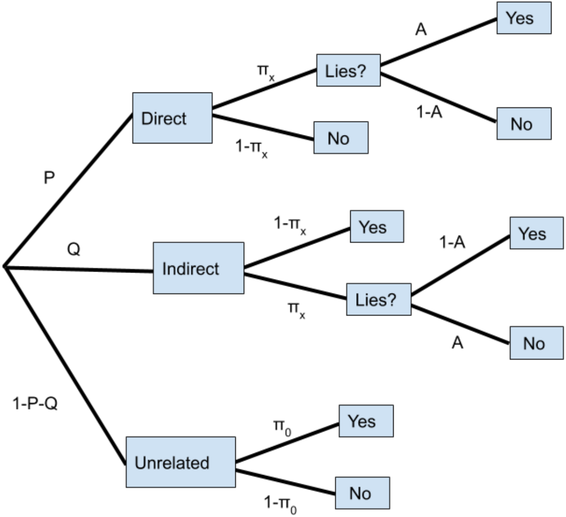
3D Renders with MatPlotLib

Turns out Warner still is more effective in alot of cases.



$n = 100, \pi_x = .3, \pi_y = .7, A = .8$

Mixed Model?



P
Q
1-P-Q
Direct
Indirect
Unrelated
π_x
1-π_x
Lies?
No
A
1-A
Yes
No
1-π_x
π_x
Yes
Lies?
1-A
A
Yes
No
π_0
1-π_0
Yes
No

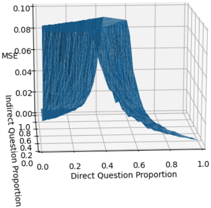
Without Incorporating Lies:

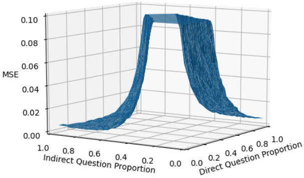
$$\hat{\pi}^*_M = \frac{P(y) - q - \pi_y(1-p-q)}{p-q}$$

With Incorporating Lies:

$$\hat{\pi}_M L^* = \frac{P(y) - qA - A\pi_y(1-p-q) - (1-A)(q + \pi_y(1-q-p))}{A(p-q)}$$

Graphs





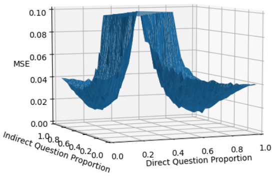
$n = 100, \pi_x = .3, \pi_y = .7, A = .8, p = 0.02, q = 0.74, MSE = .0043$

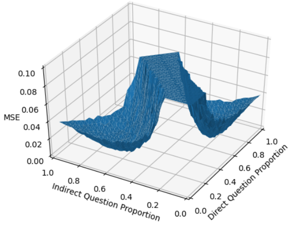
What If Lying Is A Function of the Direct Question

Could it be the case that people do not lie based if the question is sensitive but how well scrambled there response is.

For example would it be the case that the same amount of people would lie when $p = .9$ versus $p = .5$ for Greenberg or Warners model

Without Incorporating Lies

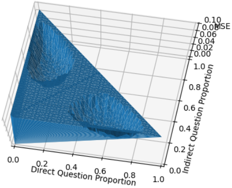


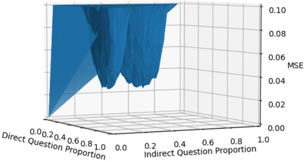


$$A = \frac{pq(1-p-q)}{(1/3)^3}$$

$$n = 100, \pi_x = .2, \pi_y = .7, p = 0.16, q = 0.62, MSE = 0.0156$$

With Incorporating Lies





$$A = \frac{pq(1-p-q)}{(1/3)^3}$$

$$n = 100, \pi_x = .2, \pi_y = .7, p = 0.12, q = 0.64, MSE = .024$$