

# Structure equation modeling

## I. OVERVIEW

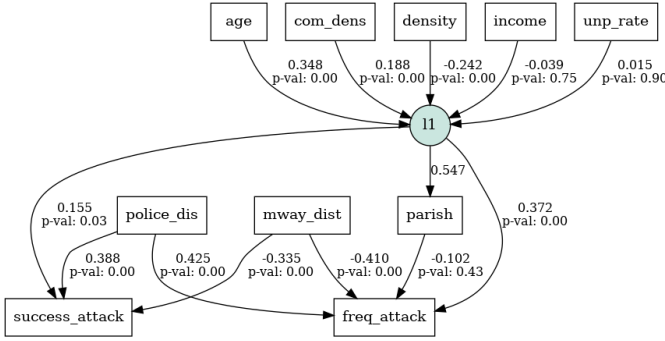


Fig. 1: SEM Model

- 1) Name of objective: Diagonally Weighted Least Squares
- 2) Optimization method: SLSQP
- 3) Objective value: 0.058
- 4) Number of iterations: 51

### Equation used

- 1)  $l1 \sim \text{parish}$  (1)
- 2)  $\text{success\_attack} \sim l1 + \text{mway\_dist} + \text{police\_dis}$  (2)
- 3)  $\text{freq\_attack} \sim l1 + \text{parish} + \text{mway\_dist} + \text{police\_dis}$  (3)
- 4)  $l1 \sim \text{com\_dens} + \text{age} + \text{income} + \text{unp\_rate} + \text{density}$  (4)

- 1) DoF: 39
- 2) chi2: 41.931951
- 3) chi p-value: 0.344902
- 4) CFI: 0.998558
- 5) GFI: 0.979888
- 6) RMSEA: 0.010204

TABLE I: Parameter Estimates

Variables	Estimate	Std. Err	z-value
parish $\rightarrow$ L	1.000000	null	null
L $\rightarrow$ com_dens	0.035499	0.011404	3.112 724
L $\rightarrow$ age	0.106035	0.020269	5.231 404
L $\rightarrow$ income	-0.009657	0.029956	-0.322 363
L $\rightarrow$ unp_rate	0.008836	0.07163	0.123 352
L $\rightarrow$ density	-0.075234	0.018266	-4.118 882
success_attack $\rightarrow$ L	0.015101	0.007025	2.149 691
success_attack $\rightarrow$ mway_dist	-0.000104	0.000012	-8.371 719
success_attack $\rightarrow$ police_dis	0.000357	0.000031	11.462 717
freq_attack $\rightarrow$ L	0.054424	0.015723	3.461 409
freq_attack $\rightarrow$ parish	-0.008179	0.010427	-0.784 397
freq_attack $\rightarrow$ mway_dist	-0.000191	0.000018	-10.629 688
freq_attack $\rightarrow$ police_dis	0.000589	0.000045	13.055 707

## II. MODEL EXPLANATION

**Magnitude of Coefficients:** The magnitudes of the coefficients indicate the strength and direction of the associations. A larger magnitude suggests a stronger association between the variables.

### A. Geographic

It was already identified with scatter plot visualization that police distance and motorway distance share negative and positive linear association relationship respectively with frequency and success of attack.

The SEM model sheds lights on the following statements: -

- 1) Positive association of police distance with success and frequency of attack indicates as the value of "police distance" increases, the success and frequency of attack tends to increase as well and vice versa.
- 2) Negative association of motorway distance with success and frequency of attack indicates as the value of motorway distance increases the success and frequency of attack decreases and vice versa.

### B. Demographic

The demographic factors do not influence the success and frequency of attack directly, they influence it through a latent variable **L**. **L** is a latent variable measured using parish to see the influence of demographic factors indirectly through **L** on success and frequency of attack.

#### Association with L

- 1) Age and Commercial density share a positive association with **L**, which can be interpreted as increase in age rate and commercial density of an area leads to increase in the latent factor **L**.
- 2) Population density is associated negatively with **L** indicating with increase in population density the latent variable tends to decrease.
- 3) Income and unemployment rate of the area have very low association in addition to that the P value ( $>0.05$ ) of these factors indicate that they are not statistically significant.

#### Association of L with success and frequency of attacks

The latent variable **L** is measured using parish which is a categorical column ranging from 1 to 20 and each point represent a local neighborhood in Lisbon.

- 1) **L** is positively associated with both frequency and success of attack with p value indicating that the association is statistically significant.

### C. Conclusion

Parish direct influence on frequency of attack is not statistically significant but the latent variable measured using parish which is predicted using demographic variables shows a positive association with frequency and success of attack.

Increase in latent variable means with increase in  $L$ , the average age rate and commercial density of the area increase and population density decreases which overall lead to increase in success and frequency of attack and vice versa.

Parishes direct influence on success of attack is not statistically significant (identified by adding parish in 2nd equation) as such that path was removed as it also caused the model behavior change (i.e. Introducing a CFI value greater then 1).

### III. ADDITIONAL INFORMATION

As our dataset is biased with more values of non attacked ATMs then attacked ATMs, The above model was run through a bias correction technique which resamples the mean of the data to adjust parameter estimates in statistical models to account for potential biases in the data or estimation procedure. This techniques aimed to improve the accuracy and reliability of parameter estimates by accounting for various sources of bias, such as sampling variability, measurement error, or model misspecification.

Resampling techniques involve repeatedly sampling from the data or model and recalculating parameter estimates to assess their variability and potential bias.

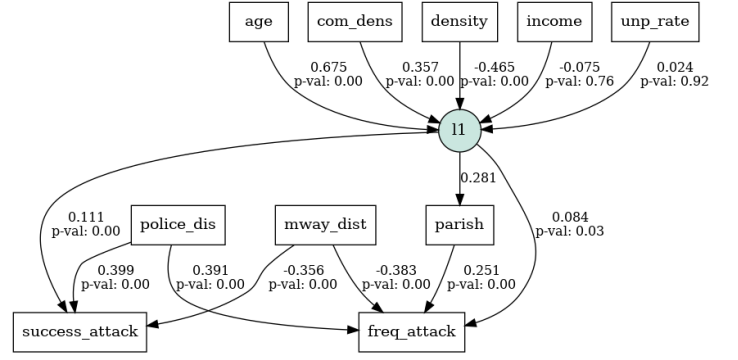
The number of resampling samples used in the bias correction process in  $n = 500$ .

For which the result obtained with resampling the mean and without resampling the mean were same. The only difference in the bias corrected model is that it shows parish has a direct influence on frequency of attack.

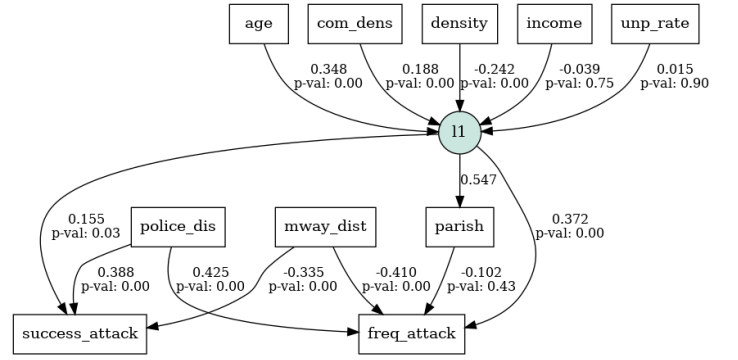
Bias correction can have a significant influence on the results of a structural equation model. By accounting for biases in parameter estimates, bias correction techniques can lead to more accurate and reliable model estimates. Bias correction helps to improve the validity and robustness of the conclusions drawn from the SEM analysis, enhancing confidence in the relationships and patterns identified in the data.

Overall, with resampling and mean correction parameters helps to address potential biases in SEM parameter estimates, leading to more reliable and robust results. It is an important step in ensuring the validity and accuracy of SEM analyses, especially in situations where data quality or model assumptions may be uncertain.

### Comparison of models



(a) Bias Corrected Model



(b) Original Model

Fig. 2: Comparison of SEM Models