ORIGINAL ARTICLE

SAMPLE SIZE CALCULATION IN STRUCTURAL EQUATION MODELING OF EQUILIBRIUM

DOI: 10.36740/WLek202312113

Oleksandr M. Ocheredko, Oleg V. Klimenyuk, Anastasya A. Rudenko, Volodimir P. Klimenyuk
PIROGOV NATIONAL MEDICAL UNIVERSITY OF UKRAINE, VYNNITSIA, UKRAINE

ABSTRACT

The aim: To pave the way and exemplify sample size calculation to studies with complex data structures describing equilibrium.

Materials and methods: Try is probated to apply ad-hoc power analysis to structural equation modeling (SEM) of equilibrium. As example we use theoretical structural equation system that describes equilibrium of price, quality and comfort of health services developed by Dranove D., Satterthwaite M. (1992). We show the way of transition from theoretical balance models to SEM that can be processed with common statistical tools. SEM is prerogative for such transition as demonstrated in the paper. We introduced some new ideas to support transition. We use Satorra & Sarris (1985) method of ad-hoc power analysis to SEM.

Results: The sample size to support error types 1 and 2 at arbitrary accepted levels 0.05 and 0.2 is 400 at least to test the influence of equilibrium price (p*) on equilibrium quality (q*). 600 sample size is needed to check for the influence of equilibrium price (p*) on equilibrium comfort (c*). Sample size of 600 is required to test hypothesis on informational noise about quality influences equilibrium value of quality.

Conclusions: It's new ground that we are tentatively exploring in paper concerning SEM of equilibrium. The main challenge as we see it is the transition from theoretical balance models to SEM that can be processed with common statistical tools. SEM is prerogative for such transition as demonstrated in the paper.

KEY WORDS: SEM, equilibrium, Power Analysis, Health Services Efficiency

Wiad Lek. 2023;76(12):2641-2647

INTRODUCTION

Sample size calculus is important to put a credit to p-value of effect. Take for example historical Tromboembolism Data. This case– control data first considered by Worcester J. (1971) [1]. The data cross-classify thromboembolism and control patients by two risk factors: oral contraceptive user and smoking. Data are regularly used to compare count data estimators, and in subsequent model choice studies, such as Spiegelhalter and Smith (1982) [2], Pettit and Young (1990) [3], Congdon P. (2005) [4], Ocheredko O (2019) [5].Under the potentially informative priors used, the Bayes factor estimate was B_{21} = 23.8, quite strongly in favour of the smaller model with single interaction effect contraceptive*thromboembolism which in all tested estimators proved to be significant. The question is should we put a credit to these findings given the original sample size of 174? To resolve the issue we have to do power analysis. This example is simple to get on but what if we have complex data structure with some variables unobservable or measured with error? What if records are structurally related (e. g., evince nesting, spatial or temporal correlation patterns)? The pick of the bunch is structural equation modeling (SEM) that usually applied to complex data that support simultaneous or consequential testing of multiple hypotheses. Particular difficulties are imposed by modeling equilibrium processes. Challenges arise at the stage of transition from theoretical structural equations that define equilibrium process (momentum or dynamic) to operational SEM formulation.

There are two general branches of power analysis, ad-hoc and post-hoc. Ad-hoc power analysis uses expert opinion on parameters comprised by statistical test, for example expected difference of two indexes and its sample error for 2 independent samples t-test. Post-hoc counterpart relies both on expert opinion and preliminary data. Actually data can be fed from accomplished study to check the relevancy of p-values.

Structural equations sometimes are substantiated by theoretical equations, which is the common case in health econometric applications. The obstacle to overcome is that theoretical equations may include unobservable variables, like survival skills, patient's utilities, satisfaction, et al. In our test case theoretical econometric equations of Dranove D., Satterthwaite M. include demand elasticities, as well as equilibrium values that are either unknown or unobservable or both.

It's very important to transform such theoretical equations into practical structure supported by statistical data. That's why we cover such transformation briefly, following with sample size calculation on resultant structural equation model.

Overall this paper focuses on power/ sample size issue that is important to support complex data describing equilibrium and in that it blazes the trail.

THE AIM

The aim of the study was to pave the way and exemplify sample size calculation to studies with complex data structures describing equilibrium.

MATERIALS AND METHODS

These includes explaining of (i) theoretical structural equations used in example, (ii) conversion of theoretical structural equations to structural equation model, (iii) power analysis technique.

Theoretical structural equations used in example are derived based on profit maximizing behavior of provider who seeks for profit maximizing levels of three attributes: price (p), quality of services (ql), and comfort (cm). Demand is function of these, i.e., q(p,ql,cm), decreasing in price and increasing in ql and cm. The total cost are function of demand q, as well as ql and cm, so that C(q, ql, cm)=q•(a+b•ql+c•cm) + F with a+b•ql+c•cm is constant marginal cost of production and F is fixed costs. So, the profit is the function of p, ql, and cm:

Profit = p•q• (p,ql,cm) - C(q(p,ql,cm), ql, cm) = q(p,ql,cm)(p - b•ql - c•cm) – F

By taking first derivatives by attributes p, ql, cm and solving for their optimal (equilibrium) values given attributes elasticities of demand $(\eta_p^q, \eta_{ql}^q, \eta_{cm}^q)$ we can conceptualize theoretical model in system of three simultaneous equations derived by Dranove D., Satterthwaite M. (1992) [6]:

$$\mathrm{p}^* = \frac{a\eta_p^q}{1+\eta_p^q+\eta_{ql}^q+\eta_{cm}^q} = \frac{(\mathrm{a}+\mathrm{b}*\mathrm{ql}^*+\mathrm{c}*\mathrm{cm}^*)\eta_p^q}{1+\eta_p^q}$$

$$\mathrm{ql}^* = \frac{a\eta_{ql}^q}{b\left(1+\eta_p^q+\eta_{ql}^q+\eta_{cm}^q\right)} = \frac{\mathrm{p}^*}{b}\frac{\eta_{ql}^q}{\eta_p^q}$$

$$\mathrm{cm}^* = \frac{a\eta_{cm}^q}{c\left(1+\eta_p^q+\eta_{ql}^q+\eta_{cm}^q\right)} = \frac{\mathrm{p}^*}{c}\frac{\eta_{cm}^q}{\eta_p^q}$$

Theoretical equations can't directly be put to practicalities (e.g., statistical regression or power analyses) because they include elasticities of demand $(\eta_p^q, \eta_{ql}^q, \eta_{cm}^q)$ that are unavailable ad hoc as well as equilibrium values of price, (p*), quality (ql*), and comfort (cm*) of services which are not directly available and that can only be elucidated by appropriately built model frame.

CONVERSION OF THEORETICAL STRUCTURAL EQUATIONS TO THE STRUCTURAL EQUATION MODEL

First of all we defined non-measurable variables with the related observables. In terms of SEM non-measurable variables are latent factors. Observables are measurable variables used to define latent factors by loadings. Latter are essentially regression coefficients so that latent factor regressed on related observables and so defined by them. For instance, quality of hospital services (F2) is latent factor defined by hospital length of stay, surgery complications risks, quality of personnel, etc. The other latent factors are comfort of hospital services (F3), and information noise (F1). Presence of F1 is a trick to render elasticities effects for F1 greatly influences all three of them $(\eta_p^q, \eta_{ql}^q, \eta_{cm}^q)$. There are no demand elasticity coefficients per se among associations. Their influence on equilibrium values of attributes is traceable through associations of F1 with price, F2, and F3. Given their key role in equilibrium related hypotheses formulation correspondent arrows rendered in red in Fig.1. Equilibrium values are substituted with observed values of attributes. It is sensible for theoretical demand and realized demand for surgeries almost coincide. Would we had have arrived at significant regression effects of F1 on price (b11 in Fig.1), quality of hospital services (z1), and comfort (z2) we would conclude that observed values of attributes are not so far from the equilibrium and so we have efficient production of health services. Supportive to theoretical equilibrium equations of Dranove D., Satterthwaite M. are regression effects of price on quality of hospital services (b31) and comfort (b31) that are colored green.

We use standard graph presentation of SEM (Fig.1). Latent factors are encircled and named with beginning letter "F", observable variables are beveled with rectangles, single-headed arrows denote directional associations, while two-headed indicate variances and covariative associations. Names of factor loadings in graphical presentation usually start with "f". Numbers assigned to arrows infer the magnitude and direction of associations. Graphical SEM model as depicted in Fig.1 is produced in special SEM tools environment accessible through https://webpower.psychstat.org/wiki/ with details given by authors [7].

Values of model's parameters (i.e., variances, covariance, regression coefficients, factor loadings) are retrieved from published sources. These and practicalities of conversion of given theoretical structural equations to the structural equation model is delivered at length in [8].

POWER ANALYSIS TECHNIQUE

Having defined SEM in graphical form like that described by Fig.1 is enough to proceed to power calculus. Given

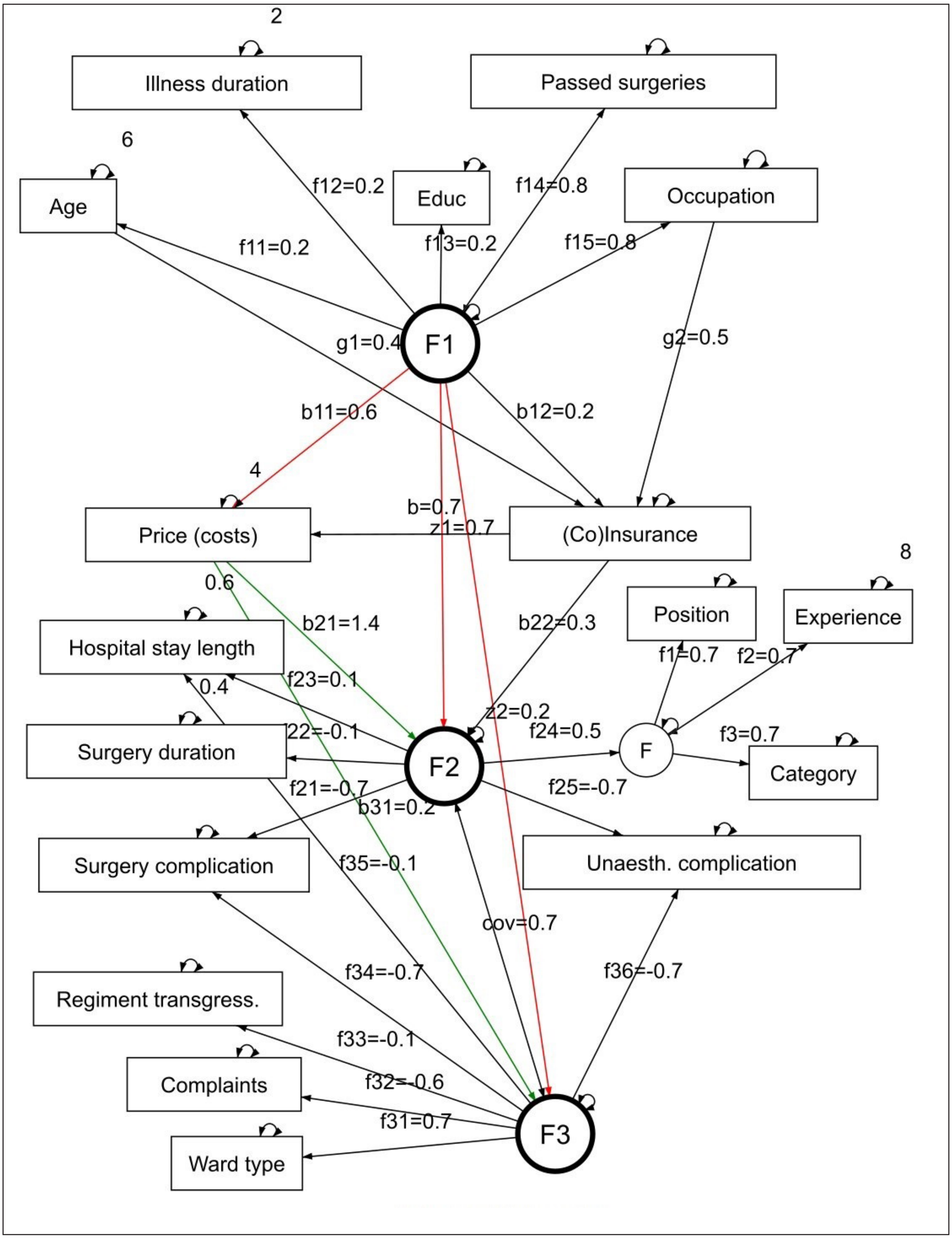

Fig.1. SEM model graph

SEM complexity and possible parameters dependency introduced by latent factors power routine should be based on statistical test that incorporate covariance matrix of parameters or technically speaking Hessian. These should be produced by routine. The commonly used under circumstances is likelihood ratio test. The pick of the bunch is the one based on chi-square test implemented with Satorra & Sarris (1985) method [9]. In brief let S denote an unbiased sample covariance matrix and θ let denote parameters in a SEM model. Let Σ be the covariance matrix defined by the model with parameters θ. From SEM theory, statistic

$$\hat{W}=(n-1)\log|\Sigma(\hat{\theta})|+tr(S\Sigma(\hat{\theta})^{-1})-\log|S|-p$$

```
full.model <-'
f1 =~ start(0.2)*age + start(0.2)*t_ill +start(0.8)*srg +
start(0.2)*educ +start(0.8)*occup
age ~~ start(6)*age
t_ill~~start(2)*t_ill
cost~~start(4)*cost

cost ~ start(0.6)*f1 + start(0.7)*insur
insur ~ start(0.2)*f1 +start(0.4)*age +start(0.5)*occup

t_stay ~~ start(0.6)*t_stay
t_surg ~~ start(0.4)*t_surg
c_surg ~~ start(1.0)*c_surg

f2 =~ start(0.1)*t_stay + start(-0.1)*t_surg +start(-0.7)*c_surg +
start(-0.7)*c_anesth  + start(0.5)*f0
f2 ~ start(0.3)*insur + start(1.4)*cost + start(0.7)*f1

expert ~~ start(8)*expert
f0 =~ start(0.7)*expert + start(0.7)*pos + start(0.7)*cat
f3 =~ start(-0.1)*t_stay + start(-0.1)*t_surg +start(-0.7)*c_surg +
start(-0.7)*c_anesth  + start(-0.1)*defect + start(-0.6)*compl + start(0.7)*ward
f3 ~ start(0.2)*cost + start(0.2)*f1
f2 ~~ start(0.7)*f3
'
```

Fig. 2. Model description in SEM language (R)

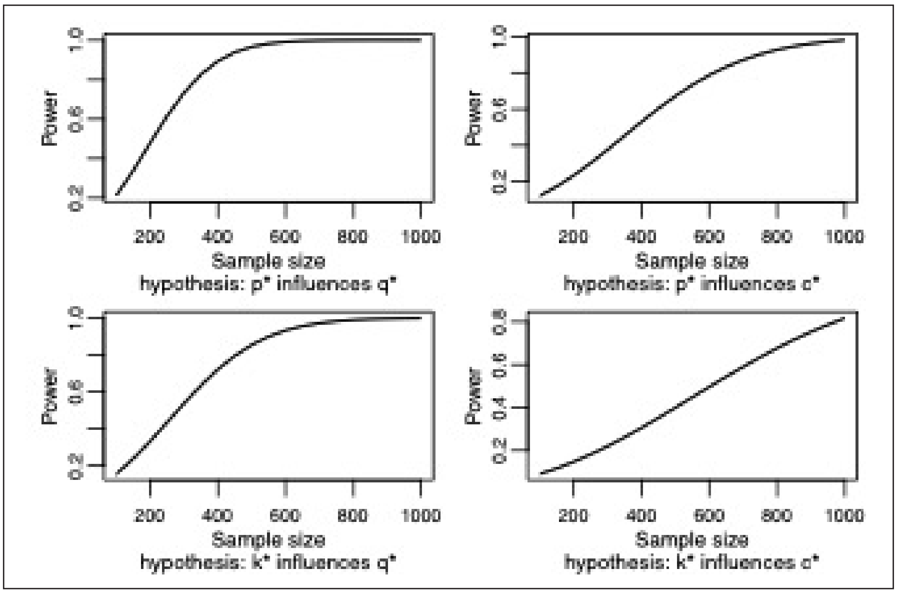

Fig.3. Power curves to test 4 hypotheses: equilibrium price (p*) influences equilibrium quality (q*) and comfort (c*), informational noise about quality and comfort influence q* and c* respectively

follows a chi-squared distribution with degrees of freedom d asymptotically. The purpose is to test the hypothesis that H_0 :θ=θ$_0$ vs H_1 : θ = θ$_1$. Under H_0, we have $P(\chi^2_d > c_\alpha) = \alpha$ where c_α is the critical value under the chi-squared distribution with degrees of freedom d. Under H_1, $\hat{W}$ follows asymptotically a non-central chi-squared distribution with the non-centrality parameter λ. The statistical power is defined as Power = $P(\hat{W} > c_\alpha | H_1)$.

Satorra & Sarris (1985) showed that λ can be approximated by

$$\lambda \approx (n-1)[\log|\hat{\Sigma}_R| + tr(\Sigma_F \hat{\Sigma}^{-1}) - \log|\Sigma_F| - p]$$

where Σ_F and Σ_R are defined under H_1 and H_0, respectively. With this, one can define an effect size independent of sample size as δ = λ/(n – 1). The effect size is defined as the difference between two SEM models, a full model M_F and a reduced model M_R. The full model (correct population model) includes all the parameters in the population (Fig.1) and the reduced model is nested within the full model by setting certain relationship to be null (H_0 hypothesis). An easy way to get the effect size is to fit the reduced model to Σ_F through SEM software with a predefined sample size n to get the chi-squared statistics λ. In depth technicalities of the approach are given in Yuan K.-H., Zhang Z., Zhao Y. [10]

We have checked four hypotheses, each H_0 hypothesis is described in SEM language first two concern price influence quality (F2) and comfort (F3). For these H_0 hypotheses are represented by removal from full model arrows defined coefficients b21 and b31 that formulates reduced models 1 and 2 (Fig. 1). Third and fourth hypotheses check for informational noise consequence on F2 and F3, so that H_0 hypotheses are represented by removal from full model arrows defined coefficients z1 and z2 (reduced models 3 and 4).

We used R package *WebPower* to do calculus. Function *sem()* is used to calculate covariance matrix, *sem.effect.size()* produces effect size, *wp.sem.chisq()* works out power reckoning, *plot()* method for class 'webpower' is used to plot the power curves.

RESULTS

We don't show covariance matrix for lack of space but calculus is reproducible given model description of full model (Fig. 2)

Code is standard for SEM and having Fig.1 is easy to read for it describes the graphical model. Values in parentheses following *start* keyword are those displayed by graph. Effect sizes estimates play key role in power analysis. They are evaluated by *sem.effect.size()* functions with two arguments which are full model, obtained by code of Fig.2 and reduced model obtained with the same code but with tested parameters put with zeroes, i.e., *start(0)*parameter_name*.

So, we retrieve effect sizes for four hypotheses by code:

```
effect.res1 <- sem.effect.size(full.model, reduced.model1)
effect.res2 <- sem.effect.size(full.model, reduced.model2)
effect.res3 <- sem.effect.size(full.model, reduced.model3)
effect.res4 <- sem.effect.size(full.model, reduced.model4)
```

Finally, we build power curve objects based on effect sizes with the help of function *wp.sem.chisq()* with arguments indicating range of sample sizes to study (n), degrees of freedom of the chi-squared test (df) and effect size (effect) both calculated by *sem.effect.size()* functions.

```
pwr.curve1 <- wp.sem.chisq(n=seq(100, 1000, 50), df=effect.res1$df, effect=effect.res1$delta, power=NULL)
pwr.curve2 <- wp.sem.chisq(n=seq(100, 1000, 50), df=effect.res2$df, effect=effect.res2$delta, power=NULL)
pwr.curve3 <- wp.sem.chisq(n=seq(100, 1000, 50), df=effect.res3$df, effect=effect.res3$delta, power=NULL)
pwr.curve4 <- wp.sem.chisq(n=seq(100, 1000, 50), df=effect.res4$df, effect=effect.res4$delta, power=NULL)
```

plot() method is used to plot the power curves, e.g., *plot(pwr.curve1)* to plot power curve for the first hypothesis.

Produced with *plot()* method power curves are demonstrated in Fig. 3.

Power curve shows relationship between sample size and power. Power value shows how reliable is p-value. Say, power 0.9 safeguards 90% reliability, so that p-value is not incidental, that is only 10% to the chance for p-value to exceed the given. Power curves are built for p-value 0.05 as regular. We follow the suit. Usually sample size is telling enough given power higher 0.8. Power curves are ascending with gradual leveling off at larger sample sizes.

From the power curves of the case we judge the sample size to support error types 1 and 2 at arbitrary accepted levels 0.05 and 0.2 is 400 at least to test the influence of equilibrium price (p*) on equilibrium quality (q*). 600 sample size is needed to check for the influence of equilibrium price (p*) on equilibrium comfort (c*). Sample size of 600 is required to test hypothesis on informational noise about quality influences equilibrium value of quality. The most required is sample size to test fourth hypothesis on informational noise about comfort influences equilibrium value of comfort, reaching 1000.

DISCUSSION

The complete absence of power analyses for SEM of equilibrium is explained in part by the culprits of transition from theoretical balance models to SEM that can be processed with common statistical tools. SEM is prerogative for such transition as demonstrated in the paper. Dynamic equilibrium can be presumably rendered by state models, changing point models, antedependence models, etc. Whatever base formulation is considered next step is necessity to wrap it into SEM because as often as not equilibrium implies contemporaneous or lagged congruous move of many variables, each move described by partial derivate. So instead of single we have set of structurally related theoretical equations to pass on to SEM. We demonstrated how we can manage transition with example. Without doubt each case is different but some common tricks are there to use. Theoretical notions like patients' preferences, propensities, idiosyncrasy in response to treatment can be rendered with latent factors which are manageable within SEM. Other constituents of theoretical equation system describing equilibrium can be settled with additional beacons that greatly influence given constituents, like factor 1 in example unfortunately of latent nature but in other cases observable. This is our first try in the matter so experience is lacking, but hopefully the growing number of researches that can be enforced by theory with equilibrium support will yield further advancements, may be quite different from suggested in the paper. As for now, some recommendation to power analysis for SEM of equilibrium can be tentatively suggested as discussion points.

First of all, ad-hoc power analysis confided to particular structure (Fig.1) with no lee-way. Information on parameters fed from other studies or expert opinions. The actual study data may not comply with ad-hoc values compromising derived sample size. Is it advisable to reassess sample size in the process of data collection? At least it is a possibility. Such power analysis is well known as ad-hoc. The new flexible approach is suggested by Ocheredko O. (2019) [5] that can be used to refine sample size estimates. It is implemented in R package *ltable* for categorical complex data and supports power analysis for simultaneous and consequential hypotheses testing.

Second, all four hypotheses should be tested jointly as ingrained in the same structure. The discussion point is what software is better for a purpose. Besides *WebPower*, other R packages can be used with even greater flexibility for ad-hoc power analysis for related tests, *lavaan* and *nimble* are paragon.

Third, given practicalities of data collection we may be unable to collect so many as power analysis suggests. It is usual situation that prompts to combine data with other sources or future data augmentation. Both options bring deduction closer to final. The point is that post-hoc power size calculus usually requires larger sample size. Therefore, ex post power evaluations of p-values are less optimistic, what was demonstrated in [5] and in manual to R package *ltable* with examples. It undermines deductions and findings of the accomplished study.

We also experience some uncertainty with Satorra & Sarris method. Caution is advisable for it compares two covariance matrices given main and zero hypotheses but it doesn't discriminate SEM structures. Several zero hypotheses with different SEM structures may result in very similar resultant matrices (discriminants/traces).

CONCLUSIONS

Power analysis for research in every field is of paramount importance. Without it we can't put credence to p-values of findings. The power analysis is of two branches, ad-hoc and post-hoc with different techniques applied. Yet it is far from perfection and there are constant debates about inconsistencies. The main is whether to get along with ad-hoc or rather with post-hoc. With ad-hoc analysis it's hard to process complex data, like that feeding to SEM. It's new ground that we are tentatively exploring in paper concerning SEM of equilibrium. The main challenge as we see it is the transition from theoretical balance models to SEM that can be processed with common statistical tools. SEM is prerogative for such transition as demonstrated in the paper.

Having hypotheses framed in solid theory gives hand to model elaboration and identification. SEM fits the purpose to a tea for its ability to incorporate structural dependencies and covariates along with directly unobservable factors. Power analysis is of particular reliability given complexity of construct that influence also possibility to test several hypotheses at once. Further post-hoc power analysis is needed to refine sample size for ad-hoc analysis can't comprise all specifications and possible measurement biases.

ABBREVIATIONS

MCMC – Monte Carlo Marcov Chain
OR\RR – odds ratio \ relative risk
SEM – structural equations modeling
r – correlation coefficient
β – regression coefficient

REFERENCES

1. Worcester J. The relative odds in the 2 by 3 contingency table. American Journal of Epidemiology. 1971;93(3): 145-149. doi: 10.1093/oxfordjournals.aje.a121240.
2. Spiegelhalter D, Smith A. Bayes factors for linear and log-linear models with vague prior information. Journal of the Royal Statistical Society, Series B. 1982;44(3): 377–387. doi:10.1111/j.2517-6161.1982.tb01217.x.
3. Pettit L, Young K. Measuring the effect of observations on Bayes factors. Biometrika. 1990;77(3): 455–466. doi:10.1093/biomet/77.3.455.
4. Congdon P. Bayesian models for categorical data. John Wiley & Sons Ltd. England. 2005, p.415.
5. Zhiyong Zhang, Ke-Hai Yuan. New Developments in Data Science and Data Analytics. 2020. doi: 10.35566/isdsa2019c5.
6. Dranove D, Satterthwaite M. Monopolistic Competition when Price and Quality are Imperfectly Observable. The RAND Journal of Economics. 1992; 23 (4): 518-534. doi: 10.2307/2555903.
7. Zhang Z, Yuan K-H. Practical Statistical Power Analysis Using Webpower and R (Eds). Granger, IN: ISDSA. 2018. doi:10.35566/power.
8. Klymenyuk OV, Ochered'ko OM, Klymenyuk VP, Rudenko AA. Analysis of the behavior of balance values of costs, quality and comfort of medical services depending on the patient's medical education. Klinichna ta profilaktychna medytsyna, 2023;2(24): 81-88. doi: 10.31612/2616-4868.2(24).2023.12.
9. Satorra A, Saris WE. Power of the likelihood ratio test in covariance structure analysis. Psychometrika. 1985; 50(1): 83–90. doi:10.1007/BF02294150.
10. Yuan K-H, Zhang Z, Zhao Y. Reliable and More Powerful Methods for Power Analysis in Structural Equation Modeling. Structural Equation Modeling. 2017;24(3): 315-330. doi:10.1080/10705511.2016.1276836.

ORCID and contributionship:
Oleksandr M. Ocheredko: 0000-0002-4792-8581 [B,C,F]
Anastasya A. Rudenko: 0000-0003-0444-1869[B,C,E]
Oleg V. Klimenyuk: 0000-0002-9764-8350[A,B,D]
Volodimir P. Klimenyuk: 0000-0003-0345-1360 [B,C,E]

Conflict of interest:
The Authors declare no conflict of interest.

CORRESPONDING AUTHOR
Oleksandr M. Ocheredko
Pirogov National Medical University of Ukraine
6 Pirogova St., 04112 Vynnitsia, Ukraine
tel: +380980017599
e-mail: Ocheredko@yahoo.com

Received: 06.10.2022
Accepted: 03.10.2023

A - Work concept and design, **B** – Data collection and analysis, **C** – Responsibility for statistical analysis, **D** – Writing the article, **E** – Critical review, **F** – Final approval of the article