Unpacking Extramarital Affairs

Hongying Tao

2023-11-21

Abstract

This report aims to explore the various factors that influence extramarital affairs and reveal the underlying motivations and behavioral patterns behind them through factor analysis and cluster analysis. The dataset used contains extensive information on the adult population, covering individuals' marital status, occupational background, education level and other socioeconomic factors. This study first used factor analysis method to explore the key factors that influence the development of extramarital affairs among individuals, and then used cluster analysis to identify different behavioral groups and patterns. The findings reveal the complex nature of extramarital affairs and the interplay of multiple social, psychological and economic factors. Through in-depth analysis, this study not only provides a new perspective for understanding individual behavior, but also provides practical suggestions for promoting the stability of marital relationships and family harmony. Despite the limitations of the data, the findings of this study provide a basis for subsequent exploration of extramarital affairs in different cultural and social contexts.

Data

This study is based on data from a 1969 survey conducted by Psychology Today, known as the Fair's Extramarital Affairs data set. The data set contains 601 observations, covering 9 variables: frequency of extramarital sex, gender, age, age at marriage, whether there are children, religiosity (1 to 5, strongly opposed to very religious), educational level (elementary school to advanced degree), occupation (Hollingshead classification), and marital satisfaction (1 to 5, very unhappy to very happy). These data provide a multidimensional perspective for in-depth analysis of extramarital affairs.

During data processing, gender was coded as "male" = 1, "female" = 0; whether or not there were children was coded as "yes" = 1, "none" = 0. These data provide a multidimensional perspective for in-depth analysis of extramarital affairs.

Factor analysis

This section explores multiple key dimensions that influence extramarital sexual behavior through factor analysis. Factor analysis is a statistical method used to identify latent structures in observed variables. Through this approach, we can understand the different factors that influence individuals to seek sexual relationships outside of marriage and try to uncover the underlying motivations behind these behaviors

1. Method

We use a data set that includes variables such as `affairs`, `gender`, `age`, `years of marriage`, `presence of children`, `religiosity`, `education`, `occupation`, and `marital rating`. Data preprocessing includes converting gender and whether you have children fields into numerical variables. In the analysis, we apply varimax rotation to optimize factor loadings.

2. Choose the number of factors

Our goal is to determine the optimal number of factors that should be retained in our factor analytic model by evaluating the proportion of variance each factor uniquely contributes.

Determine the Range of Factor Numbers

For each model, we calculated `Phi` values, a measure that reflects the proportion of variance uniquely explained by each factor. Subsequently, we calculated the difference between consecutive `Phi` values, denoted as `Phi_diff`.

Phi	0.2310365	0.3947121	0.4759819	0.5425746	0.6009117

Phi_diff	0.1636757	0.0812697	0.0665927	0.0583371

We would like to choose m large enough so that Φ_m is relatively large, say ≥ 0.70 and $\Phi_{m+1} - \Phi_m$ is relatively small, say $\leq 1/9 = 0.111$. First note that changes in Φ_m become small once $m \geq 2$ but all Φ_m does not greater than 0.70. Hence, it appears that we should consider a model with the number of factors between 2 and 5.

Statistical Testing for Factor Number

```
##    Factors         PVAL
## 1        2 4.245094e-16
## 2        3 2.275249e-06
## 3        4 1.069176e-02
## 4        5 1.867185e-01
```

The p-value for the model with 2 factors and 3 factors are very small but the others could be viewed as acceptable.

Test the commonality of variables

```
##              X1...ea.fa4.uni X1...ea.fa5.uni
## affairs            0.5340434       0.5851253
## gender             0.8909397       0.6022083
## age                0.6471199       0.7027279
## yearsmarried       0.9950000       0.9388588
## children           0.3708821       0.9950000
## religiousness      0.1188844       0.1216575
## education          0.3859648       0.7811229
## occupation         0.7617333       0.4904027
## rating             0.1786037       0.1911019
```

Several variables, including `children`, `religiousness`, `education`, and `rating`, exhibit communalities below 0.40 in the 4 factors model. This suggests that these variables are inadequately explained by the model comprising 4 factors.

In the 5 factors model, however, certain variables such as `religiousness` and `rating` continue to show communalities below 0.40. Despite this, they remain poorly explained by the model.

Taking the above analysis results into consideration, we decided to use a model with 5 factors for further analysis. This model captures more information and explains the structure in the data better than the 4 factors model.

By choosing a model with 5 factors, we can gain a more complete understanding of the observed data and conduct a deeper analysis.

3. Results

```
##
## Call:
## factanal(x = ea, factors = 5, rotation = "varimax")
##
## Uniquenesses:
##       affairs         gender          age  yearsmarried       children
##         0.415          0.398        0.297         0.061          0.005
## religiousness      education   occupation        rating
##         0.878          0.219        0.510         0.809
##
## Loadings:
##                Factor1 Factor2 Factor3 Factor4 Factor5
## affairs                                        0.761
## gender                                 0.742
## age            0.790
## yearsmarried   0.896
## children                       0.923
## religiousness
## education              0.830
## occupation             0.472           0.497
## rating
##
##                Factor1 Factor2 Factor3 Factor4 Factor5
## SS loadings     1.662   0.997   0.960   0.917   0.872
## Proportion Var  0.185   0.111   0.107   0.102   0.097
## Cumulative Var  0.185   0.295   0.402   0.504   0.601
##
## Test of the hypothesis that 5 factors are sufficient.
## The chi square statistic is 1.74 on 1 degree of freedom.
## The p-value is 0.187
```

Factor 1: "Marriage Durability" factor is highly correlated with age and years of marriage, which may reflect that as age increases and the duration of marriage increases, individual behavior in marriage may change.

Factor 2: "Education and Occupation" factor, the high loadings of education level and occupational status indicate the influence of these socioeconomic factors on extramarital sex.

Factor 3: "Children and Family" factor, the high loading of the presence of children indicates that family structure and children may be key factors affecting extramarital sex.

Factor 4: "Gender role" factor, the high loading of gender may point to the impact of gender roles and gender expectations on individual behavior

Factor 5: The "frequency of extramarital sex" factor is directly related to the frequency of extramarital sex, which may represent the sum of other factors that directly affect the frequency with which individuals seek sexual relationships outside of marriage.

Model Fit

In factor analysis, the model fit was primarily based on the chi-square test results. According to the analysis output, the p-value of the chi-square test is 0.187. This result is higher than the threshold of 0.05 usually used for significance judgments, meaning that we do not have sufficient statistical evidence to reject the hypothesis of a good model fit. In other words, this P value indicates that from a statistical perspective, our model fits the collected data well. This indicates that, given the number of variables and factors considered, the model is relatively effective at explaining the observed correlations between variables in the data.

4. Summary

This factor analysis summary reveals five key factors that influence extramarital sex. First, the "marriage durability" factor suggests that an individual's behavior in a marriage may change as they age and have been married for longer. Secondly, the "education and occupation" factor emphasizes the impact of educational level and occupational status as socioeconomic factors. Third, the "Children and Family" factor highlights the importance of family structure and children. Fourth, the "gender role" factor reveals the impact of gender roles and expectations on individual behavior. Finally, the "frequency of extramarital sex" factor is directly related to the frequency with which individuals seek sexual relations outside of marriage, reflecting the combined influence of other factors. Together, these factors form a complex framework that influences individual behavior in marriage.

Cluster Analysis

This section uses cluster analysis to analyze the data and aims to identify different possible patterns of infidelity. This data set contains multiple variables related to individuals' marital status and personal behavior, such as `affairs`, `gender`, `age`, `years of marriage`, `presence of children`, `religiosity`, `education`, `occupation`, and `marital rating`.

1. Method

The analysis utilizes the K-means clustering algorithm to separate survey data into different groups based on variables reflecting marital status and personal behavior. The optimal number of clusters was determined and the random initialization of the algorithm was controlled by setting a reproducibility seed. This methodology underpins our goal of discerning patterns that may indicate a propensity for extramarital affairs.

2. Result

affairs	gender	age	yearsmarried	children	religiousness	education	occupation	rating
0.7534247	0.4246575	24.13699	2.393470	0.3789954	2.835616	15.88128	4.091324	4.255708
1.9285714	0.5659341	43.89560	14.631868	0.9505495	3.489011	16.42857	4.538462	3.620879
1.7950000	0.4500000	31.25000	8.638125	0.8700000	3.085000	16.24000	3.995000	3.860000

It is obvious that the average number of affairs among individuals in cluster 2 is 1.93, which is much higher than the other two clusters. This suggests that individuals within the cluster are more likely to engage in extramarital affairs.

Let's analyze in detail the salient characteristics displayed by the individuals in cluster 2:

Gender distribution: The average code is 0.566, implying that males are the majority in this cluster.

Age: The average age of individuals in the cluster is 44 , concentrated in the middle age group.

Years of marriage: The average age at marriage for individuals is 11.6 years.

Children: Most individuals have children.

Level of religiosity: The average level of religiosity is 3.5, indicating a low level of religiosity.

Education level: Highly educated, average education level is college graduate.

Occupational status: The average occupational level is 4.5, and most are in professional fields or management positions.

Marital satisfaction: The average marital satisfaction score is 3.6, which is lower than the average. Individuals are relatively less satisfied with their marriage.

3. Summary

In this part, it can be seen from the characteristics of cluster 2 that individuals in this group may face lower marital satisfaction, which may be related to their higher tendency to have affairs. These individuals are typically middle-aged men with children, high levels of education, and stable careers. They may be dissatisfied in their marriage due to the pressure and responsibilities of a professional or managerial position. At the same time, lower levels of religiosity may be associated with weaker associations with traditional marital values. These findings suggest that marital quality may be a key factor influencing extramarital affairs. This situation reflects how complex social and psychological factors come together to influence marital relationships and individual behavior.

Conclusion

Through careful analysis of the *Fair's Extramarital Affairs* data set, this study reveals a variety of key factors that influence extramarital affairs. Our analysis highlights the role of factors such as age, age at marriage, gender, presence of children, religious affiliation, education level, occupation, and marital satisfaction in extramarital affairs. It is particularly worth noting that the factors of gender and having children show a significant correlation in the occurrence of extramarital affairs: men and individuals without children have a relatively higher incidence of extramarital affairs.

In addition, the frequency of extramarital affairs showed a significant negative correlation with marital satisfaction, suggesting that declining marital quality may be an important factor driving extramarital sex. At the same time, the intensity of religious belief also has a significant impact on extramarital affairs. Stronger religious beliefs seem to be able to inhibit the occurrence of extramarital affairs to a certain extent.

The impact of education level and occupation on extramarital affairs is complex. Individuals with higher education levels and professional and occupational backgrounds do not show an obvious unified trend in the occurrence of extramarital affairs. This may reflect the differences in individual behavioral choices under different socioeconomic backgrounds.

This study provides a deeper understanding of the social and psychological drivers of extramarital affairs. These findings not only provide important reference information for psychologists, sociologists, and marriage counselors, but also provide practical guidance for individuals seeking to improve the quality of their marriages and promote family stability. Although our analysis is limited by the nature of the data set itself and the survey methodology, these results still provide a valuable foundation for future more in-depth research on extramarital affairs in different cultural and social contexts.

Source

Infidelity data, known as Fair's Affairs. Cross-section data from a survey conducted by Psychology Today in 1969.

https://vincentarelbundock.github.io/Rdatasets/doc/AER/Affairs.html https://vincentarelbundock.github.io/Rdatasets/csv/AER/Affairs.csv