Journal of Business Studies Quarterly   jbsq.org

An Econometric Analysis of Insured Characteristics Impact on Group Health Insurance Loss: Case Study of Iran

Iman Ghalehkhondabi
Department of Industrial & Systems Engineering, Ohio University
E-mail: ig060113@ohio.edu

Abstract

The Iranian health insurance coverage had been a challenging part of the social welfare for decades. Some of the most important parts of these challenges return to the shortage of sufficient scientific works in the field. In this paper, the manner of health insured is studied in different provinces of Iran. The applied data are gathered from the issued health insurance policies of Asia Insurance Company, which is the biggest private insurance company of the country. The popularity of this policy is analyzed according to criteria such as geographical dispersion and education of each province. Also, the loss ratio variation of health insurance policies is studied according to insured characteristics such as age, gender and demographic characteristics. In each part, some analytic results are provided due to the authors' personal perception of Iranian society. These results can be used to present business development models in insurance companies.

Keywords: Group Health Insurance, Econometrics, General Linear Model, Iran

1. Introduction

The insufficiency of health care issues in the Iranian society had been an important place of argues for all recent governments. This challenge is in both universal access and the quality of provided services by the government and private companies. Due to the low knowledge based market, the Iranian insurance companies could not perform some valuable academic studies to gain a new knowledge about their insured. In recent years, some private insurance companies have decided to study the manner of their health insurance customers.

The importance of academic studies in insurance industry is understood in many countries for decades and many valuable studies are performed in this field. There are many aspects of insurance operations, which are studied by researchers such as equity, children's insurance, loss estimation, premium sharing, etc.:

A study by Feldman and Dowd [1] tried to estimate the welfare loss of excess health insurance. They calculated the welfare loss of surplus insurance by using the Rand data on consumer risk-aversion and elasticity. Considering insurance affects the gross price of medical care, they showed that in 1984, welfare loss was in a range of 33.4 to 109.3 billion dollars.

The continuity of children's health insurance is studied in a valuable article by Fairbrother et al. [2]. They surveyed the impact of parental job loss on children's coverage and declared that a much more effort must be done to enroll children in public insurance programs regardless of their parent job status. They showed that the children in families with an earning of less than 200 percent of the Federal poverty level mostly face to the risk of losing health coverage by their parent's job losing.

A different criterion in insurance decisions refers to distance role for immigrants, which is studied by Henry Brown [3]. He analyzed the Mexican immigrants' decisions encountering their health issues. His study shows the effective role of distance in the Mexican immigrants decisions, where immigrants from the south of Mexico are more likely to seek care in their hometowns. Although, distance is the main effective criteria in his study, but it is obvious there is a hidden tradeoff between premiums and traveling costs. This hidden tradeoff impact can be seen in immigrants' decisions.

Another important issue about the premium payments returns to the share of employee and the employer in group policies, where the employer pays a part of the premium. In such policies, increasing premium can result in different decisions by insured. Goldman, Robalino and Leibowitz [4] determined the sensitivity of employees' health insurance decisions when the costs of healthcare grow with a fast pace. The results show that single employees prefer to react to premium increases by dropping coverage, but families prefer to switch to another plan. The findings show that continual premium increases could bring on substantial increases in the number of individuals with no insurance coverage. Same studies in the field of cost and premium sharing are done by Cutler [5], Crimmel [6], Vistnes et al. [7], Sommers [8], Vistnes and Selden [9] and Baicker and Chandra [10].

Many characteristics would be considered in pricing health insurance which could make a complex tradeoff in the process of pricing. Frank and Lamiraud [11] considered some factors in pricing a health insurance including gender, age, health condition, income, job status and demographical conditions. Equity of the premium in Ghanaian insurance programs is studied by Amporfu [12]. It is shown that premiums are dissimilar in different cities of the country and

members with tertiary education pay a lower premium in comparison with those without tertiary education. The study also shows that the premium is probable to force catastrophic costs on a minority of the poor.

Pricing is one of the most important related issues of health insurance due to its effect on both consumer price index and company service level and loss payments. The actuarial pricing techniques are different in various countries around the world. A related research to our current study is done by Ekman [13] to analyze the role of health insurance on catastrophic health payments in Zambia. The econometric multivariate analysis is applied in this study to asses the factors of catastrophic health payments. The main result is that the true impact of health insurance depends on factors such as service provision oversight and quality assurance. It is emphasized that health insurance does not afford financial protection against catastrophic expenditures.

In current paper, linear regression and the econometric concepts are used to analyze the loss ratio variations for each policy due to policies human characteristics. General Linear Model (GLM) was used in early 1990s as the main method for predicting each policy condition. In recent decade, some companies used the General Additive Model (GAM) to predict the policies loss ratio. The main difference is in replacing explanatory variables and parameters in GLM with some complex functions. But, GAM has some limitations such as interaction of explanatory factors, collinear, over fitting and being restricted by parametric assumptions [14]. Having these limitations and some insufficiencies in informatics infrastructures in Iranian insurance companies, it is more useful to apply simple linear prediction instead of GAM. The reminder of current study is as follows: in section2 the popularity of private health insurance policy in Iranian provinces is studied, section 3 is dedicated to investigate specific trends in the loss ratio against insured characteristics by econometric software Eviews6. Some conclusions and comments for future research are presented in section4.

2. Health insurance popularity in different provinces of Iran

Health care expenditure is one of the most catastrophic costs of Iranian families. Definitely, in families who live in rural places or in provinces with lower levels of income, this issue gets more importance. The demographical study of issuance volume in insurance companies and loss ratio of the policies is essential to have a better view of future service level and financial transactions.

In Iran, health insurance is just sold to groups with more than 50 individuals due to a common perception of large numbers law in the insurance market. This kind of issuance forces the policies to obey the law of large numbers, which is the statistical base of all insurance policies. But, there may be some policies with lower individuals due to their portfolio in other insurance branches. Another important issue is about the validity period of health insurances, which is a year and it could be renewed each year. In our study, there exist about 2200 policies which were issued in 2011 and credible until 2012. According to the time consuming process of surveying all these policies, it is needed to perform the analyses on a sample of issued policies. Therefore, a random sample of 330 policies is selected due to Krejcie and Morgan method [15]. The selected policies are expired and their loss is completely settled by the company. This time period is selected to be sure that there is no continuing policy and all data are constant with no change.

2.1. Geographical dispersion

Iran has 31 provinces due to the last territorial division made by Cabinet in 2010. These provinces are shown in Figure 1. Iran is geographically a wide country, which make it a 4 season country in each time of the year. In order to have precise geographical analyses we have decided to study the geographical manner of insured in 5 categories as follow:

1- Central Provinces: Tehran, Alborz, Qom, Isfahan, Markazi
2- South Eastern Provinces: South Khorasan, Yazd, Kerman, Sistan and Baluchestan, Hormozgan
3- South Western Provinces: Fars, Bushehr, Kohgiluyeh and Boyerahmad, Chaharmahal and Bakhtiari, Khuzestan, Lorestan, Ilam
4- North Eastern Provinces: Razavi Khorasan, North Khorasan, Semnan, Golestan, Mazandaran
5- North Western Provinces: Hamedan, Kermanshah, Qazvin, Zanjan, Kordestan, Gilan, Ardabil, East Azerbaijan, West Azerbaijan

Figure 1 - Iran Provinces due to 2010 territorial division

As shown in Table 1, although the first category just has 5 provinces and other categories have more, more than 50 percent of selected policies are issued in central provinces. This could be related to the group issuance of health policies, so there would be a more demand by companies and industrial plants. According to political and economical importance of central provinces, it is logical to have more companies in these provinces with employers who prefer to provide a social welfare for their personnel and employees. This happens due to the fact that employees in central provinces know about their social rights and want health insurance as a part of their job contract.

Table 1 - Frequency of health insurance policies in Iran Provinces

Geographical Category	Number of Policies	Policies Percentage	Number of Insured	Insured Percentage
1	47	52.2	9921	58.3

2	9	10	2635	15.4
3	10	11.2	1081	6.4
4	12	13.3	1629	9.5
5	12	13.3	1767	10.4

There is another difference between the number of issuance in Northern provinces and Southern ones. This difference can be referred to the water shortage and climate problems in south of Iran. As it is obvious form Figure 1, Southern provinces are so wide in comparison to Northern ones, according to the fact southern states are not so appropriate for living and despite of wide territories they do not include a high population. Consequently, there was no economic boom in Southern provinces and fewer companies are activated in these areas. As a result, both employers and employees prefer to save the health insurance premiums in their pockets. On the other hand, Northern provinces are main tourist destinations for a high portion of the country and they have a long land border with neighbor countries, which makes an economic advantage to have higher level of social welfare in contrast to Southern Provinces.

2.2. Literacy Levels

The insured ratio is so low in Iranian community, maybe it refers to the low individual perceptions of risk management and the performance of insurance industry. In first view, it seems that level of social education may have an influence on the number of insured and the manner of loss settlement of each policy. Therefore, the level of literacy is considered as a criterion of influence on insurance manner. In order to match the number of educational categories with geographical ones, here the Iranian provinces are grouped to five categories. In this grouping, people who are older than 15 and can read and write are considered as literate. It is notable, this information returns to the time when Tehran and Alborz formed a same province and this grouping is performed by 30 provinces. Iran provinces are ranked by the criteria of literacy portion as follows:

1- The most literate Provinces: Tehran and Alborz, Semnan, Isfahan, Fars, Yazd, Mazandaran
2- Qazvin, Bushehr, Qom, Razavi Khorasan, Gilan, Markazi
3- Hamedan, Kermanshah, Khuzestan, Ilam, East Azarbaijan, Chaharmahal and Bakhtiari
4- Zanjan, Golestan, Kerman, Lorestan, Kohgiluyeh and Boyerahmad, Ardabil

5- Hormozgan, South Khorasan, North Khorasan, Kordestan, West Azarbaijan, Sistan and Baluchestan

As shown in Figure 2, it is an obvious correlation between literacy and issuance of health insurance. This correlation can be referred to the positive relation between literacy and being employed in an economical or industrial plant. Companies prefer to establish their plants in provinces where educated human resources are available. According to the commitment of insurance companies to issue health insurances for groups, it is logical that more issuances are occurred in more literate provinces.

Figure 2 – The insured dispersion according to province Literacy ratio
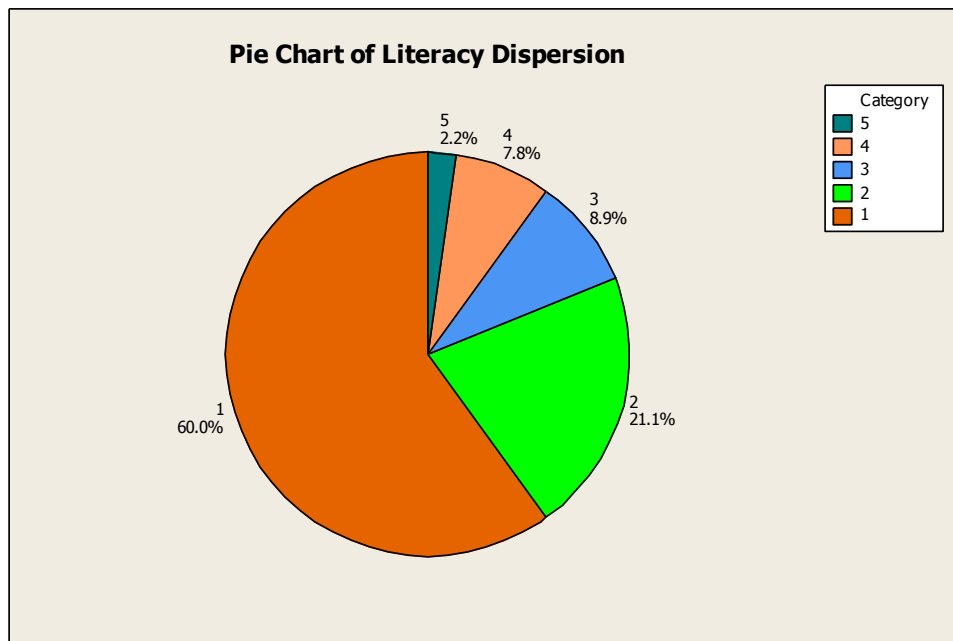
The other reason, which can be used to explain Figure 2 is the desires of literate and illiterate employees. Literate people know more about their social rights and think they must have a better job condition due to their knowledge. On the other hand, illiterate people think paying health insurance premiums is a kind of money loss and they may prefer to not have any kind of insurance.

3. Loss ratio manner against insured characteristics

As mentioned in section1, there are some studies which focused on pricing and loss ratio estimation in health insurance policies ([11], [13], [14]), but such a study is not performed in Iranian insurance market yet. An important issue about Iranian insurance policies returns to the group issuance, which fades out the characteristics of each individual in the policy. More

explicitly, the age or gender of each individual is not considered as a standard for getting the premiums. The premiums are announced by the central insurance of government due to some old actuarial tables, and companies will give some discounts, if the insured company brings an insurance portfolio beside its health insurance request. In conclusion, the characteristics of insured are reflected in policy characteristics. In our study some characteristics such as average of insured age, percentage of men, and number of insured in each policy, province of the insured company and the literacy of that province are considered as probable effective factors on loss ratio variations.

3.1. Effect of insured gender on loss ratio

There are many differences in male and female human beings, which can be effective in their healthcare issues. These differences can be reflected in different diseases and different health costs, which may influence the loss ratio of a health insurance policy. Some companies may think women are weaker than men and as a result, the policy which has more women may have a higher loss ratio. Here, the portion of men is obtained for each of 330 policies and the regression of loss ratio is estimated against men portion with software Eviews 6. Table 2 shows the results of this regression:

Table 2 – Estimation of loss ratio against gender

Variable	Coefficient	P-Value
Constant	0.517	0.238
Men Portion	0.347	0.684

As it is shown in Table 2, the variation of loss ratio is best fitted by following equation:

$$\text{Loss ratio} = 0.517 + (0.347*\text{Men portion})$$

But, it is notable the R-squared value is 0.001 which shows the inefficiency of the obtained regression model and the P-Value of both exams are more than 0.05. These P-Values show the lack of influence of a policy dominant gender on the variation of loss ratio. More explicitly, in confidence level of 95% the insurance company cannot use the policy dominant gender as a criterion to estimate the loss ratio.

3.2. Effect of insured age on loss ratio

There is a true belief about the individual life insurances that get a higher premium from older people. But this belief can be misunderstood for group insurances. This hypothesize is examined under a regression function to estimate the loss ratio against various values of mean age of each 330 policies.

Table 3 – Estimation of loss ratio against age

Variable	Coefficient	P-Value
Constant	0.548	0.147
Mean Age	0.005	0.695

Table 3 shows the output of Eviews6 for finding the regression function of loss ratio against average age of each policy. Table 3 results to the following equation:

$$\text{Loss ratio} = 0.548 + (0.005*\text{Mean Age})$$

The R-squared value for this estimation is 0.001, which means this equation predict 1thousandth of the loss ratio variations. In addition, P-Values do not show a specific impact on loss ratio. By Table 3 results, it is obvious that average age of insured in one group health policy cannot be considered as a good criterion for predicting the loss ratio.

3.3. Effect of insured number in each policy on loss ratio

The most useful statistical law in insurance industry is the law of large numbers. According to the law, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed [16]. This law forces the insurance companies to issue as more policies as they can to get closer to the expected value of the whole society and make a more precise estimation of future losses. But, this law is about the number of policies, and the effect of more individuals in a group health insurance can be different.

Table 4 – Estimation of loss ratio against number of individuals in each policy

Variable	Coefficient	P-Value
Constant	0.6037	0
Number of individuals	0.0005	0.022

Obtained P-Values in Table 4 are significant in the significance level of 95%. As well, the R-Squared value is 0.06, which is higher than obtained values in sections 3.1 and 3.2. Although, there is a low influence of number of policy individuals on loss ratio, but its effect is significant. The equation is as follows:

Loss ratio = 0.6037 + (0.0005*Number of policy individuals)

While it is not appropriate to charge a higher premium with a coefficient of 0.0005, the constant value of 0.6037 is a more reasonable value than the values obtained in sections 3.1 and 3.2. The P-Value of 0 for testing the effectiveness of constant value, guides the insurance company to consider a loss ratio of 0.6037 for offering premiums to new applicants of group health insurance.

3.4. Effect of geographical dispersion on loss ratio

It is logical to have different life styles and different levels of health care in various provinces of a geographically wide country such as Iran. As mentioned in section 2.1, in this study the country is divided to 5 categories. In order to analyze the affect of categorized explicative variables on a dependent variable, it is needed to define some dummy variables. It is possible to assign a number to each category, but this numbering is equal to assigning weights to variables, which are not proved scientifically. Therefore, dummy variables are used in such situations. It is notable, for a variable with *n* categories, *(n-1)* dummy variables must be defined [17]. Here, 4 dummy variables are defined to model the manner of loss ratio against 5 categories of provinces in Iran.

Table 5 – Estimation of loss ratio against geographical dispersion

Variable	Coefficient	P-Value
Constant	0.673	0
South Eastern Provinces	0.268	0.081
South Western Provinces	0.062	0.653
North Eastern Provinces	-0.138	0.302
North Western Provinces	0.043	0.737

The obtained equation from Table 5 is as follows:

Loss ratio = 0.673 + (0.268*SEP + 0.062*SWP − 0.138*NEP + 0.043*NWP)

The P-Value of constant coefficient is significant, but the R-Squared value of this regression is 0.05 which is lower than the obtained value of Table 4. Therefore, the insurance company may prefer to use the obtained value of Table 4 due to higher R-Squared value.

Except Constant value, none of the P-Values is significant in confidence level of 95%, but if we consider the confidence level of 90%, South Eastern provinces will have a significant effect on loss ratio. The main reason of such an effect could be the settlement of these provinces in desert of Iran. Water shortage had been the main problem in these provinces for decades and fruit consumption is so low in these provinces. According to these reasons, the company should offer higher premiums for issuing policies in South Eastern provinces.

3.5. Effect of literacy level on loss ratio

Considering insurance as a financial instrument for risk sharing, it may be logical to expect a trend in using of this instrument between literate and illiterate people. On the other hand, this policy is issued for groups where the employer is more probable to be a university graduated and employees may be simple illiterate labor, which may change the pattern of loss ratio against different levels of literacy.

Table 6 - Estimation of loss ratio against different levels of literacy

Variable	Coefficient	P-Value
Constant	0.69	0
Group2	-0.096	0.378
Group3	0.173	0.255
Group4	0.195	0.259
Group5	-0.234	0.417

The estimated regression by Table 6 is as follows:

$$\text{Loss ratio} = 0.69 + (-0.096*G2 + 0.173*G3 + 0.195*G4 - 0.234*G5)$$

Despite the constant coefficient, there is no other significant effect by other dependent coefficients. The R-squared value is 0.05, which refers to low contribution of literacy in predicting loss ratio variations. Maybe, this lack of effect returns to the tradeoff between different standards of life. In most literate provinces, health issues are more observed than less

literate provinces and this may decrease the healthcare expenditures. On the other hand, literate provinces of Iran include the biggest industrial and polluted cities and this may encounter their citizens with a high risk of disease. The tradeoff between health concerns and living environment prevents of formation a pattern in loss ratio for different levels of literacy.

3.6. Joint effect of geographical dispersion and number of policy individuals on loss ratio

The analyses in previous sections did not reflect a significant trend in loss ratio against the studied criterions. As a result, authors performed more analyses to find a trend in loss ratio against more than one factor. All the possible combinations are studied and the highest R-Squared value is related to the combination of geographical dispersion and number of individuals in each policy.

Table 7 - Estimation of loss ratio against geographical dispersion and number of policy individuals

Variable	Coefficient	P-Value
Constant	0.56	0
Number of individuals	0.0002	0.04
South Eastern Provinces	0.487	0.02
South Western Provinces	0.184	0.388
North Eastern Provinces	-0.226	0.249
North Western Provinces	-0.023	0.892
Number of individuals * South Eastern Provinces	-0.001	0.147
Number of individuals * South Western Provinces	-0.0006	0.675
Number of individuals * North Eastern Provinces	0.0008	0.374
Number of individuals * North Western Provinces	0.0006	0.377

The R-squared of this regression is 0.16 which is more than all the previous estimations. As it is inferred from Table 7, the most effective factors on loss ratio are the geographical dispersion and the number of insured in each policy. Therefore, the insurance company may charge a higher premium for policies with higher number of insured and issued in South Eastern provinces of Iran.

4. Conclusions

Many insurance companies are eager to predict the loss ratio variations against various characteristics of insured to get the appropriate premiums, which could afford the loss settlement. The loss ratio of group health insurance policies cannot be calculated based on the individuals' characteristics. In this study, the variation of Iranian health insurance policies' loss ratio is analyzed against the human and demographic characteristics of each policy. The manner of studied sample policies shows no significant effect of average age, dominant gender and literacy level of each policy on loss ratio. But, the increase in number of insured individuals in each policy and issuance a policy in south eastern provinces of Iran may increase the loss ratio. As a future research, doing a study on each individual health care costs may help the Iranian insurance companies to issue individual health insurance policies. The effect of economic situation such as inflation on health insurance costs for both insured and insurer may be an interesting field of study, as well. This study surveyed the whole loss manner for a policy, but as a future study, surveying the loss variations for each kind of coverage in a health insurance policy can get a better knowledge about the health care costs of the country.

References

[1] Feldman R, Dowd B: A New Estimate of the Welfare Loss of Excess Health Insurance. The American Economic Review 1991; Vol. 81, No. 1: 297-301.

[2] Fairbrother GL, Carle AC, Cassedy A, et al: The Impact Of Parental Job Loss On Children's Health Insurance Coverage. *Health Affairs 2010*; 29, no.7: 1343-1349.

[3] Brown HS: Do Mexican immigrants substitute health care in Mexico for health insurance in the United Provinces? The role of distance. Social Science & Medicine 2008; 67: 2036–2042

[4] Goldman D, Leibowitz A, Robalino D: Employee Responses to Health Insurance Premium Increases. The American Journal of Managed Care 2004; Vol. 10, No. 1: 41-47.

[5] Cutler DM: Employee Costs and the Decline in Health Insurance coverage. Forum for Health Economics & Policy 2003; Vol. 6: (Frontiers in Health Policy Research), Article 3.

[6] Crimmel BL: Employee Choice in Employer-Sponsored Health Insurance Plans: 2001. Center for Financing, Access, and Cost Trends, AHRQ, 2003; Medical Expenditure Panel Survey- Insurance Component: Statistical Brief #29

[7] Vistnes J, Morrisey MA, Jensen GA: Employer choices of family premium sharing. International Journal of Health Care Finance and Economics 2006; Vol. 6: 25–47.

[8] Sommers B: Who really pays for health insurance? The incidence of employer provided health insurance with sticky nominal wages. International Journal of Health Care Finance and Economics 2005; 5: 89–118.

[9] Vistnes J, Selden T: Premium growth and its effect on employer sponsored insurance. International Journal of Health Care Finance and Economics 2011; Vol. 11: 55–81.

[10] Baicker K, Chandra A: The labor market effects of rising health insurance premiums. Journal of Labor Economics 2006; Vol. 24(3): 609-634.

[11] Frank RG, Lamiraud K: Choice, price competition and complexity in markets for health insurance. Journal of Economic Behavior & Organization, Elsevier 2009; vol. 71(2): 550-562.

[12] Amporfu E: Equity of the premium of the Ghanaian national health insurance scheme and the implications for achieving universal coverage. International Journal for Equity in Health 2013, 12:4.

[13] Ekman B: Catastrophic health payments and health insurance: Some counterintuitive evidence from one low-income country. Health Policy 2007; 83: 304–313.

[14] Brockett PL, Golden LL, Jang J, et al: A comparison of neural network, statistical methods, and variable. Journal of Risk and Insurance 2006; 73(3): 397-419.

[15] Krejcie RV, Morgan DW: Determining Sample Size for Research Activities. Educational and Psychological Measurement 1970; 30: 607-610.

[16] Sheldon R: A first course in probability (8th ed.). Prentice Hall press 2009.

[17] Salvatore D: Statistics and Econometric. Mcgraw-Hill 1981.