

## Applying Multiple Linear Regression to Predict Women Representation

<sup>1,2</sup>Najeebullah, <sup>3</sup>Adnan Hussein and <sup>4</sup>Bakhtiar Khan

<sup>1</sup>School of Communication, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia

<sup>2</sup>Department of Public Administration, Gomal University, Dera Ismail Khan and Pakistan

<sup>3</sup>School of Communication, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia

<sup>4</sup>Department of Business Administration, Gomal University, Dera Ismail Khan, Pakistan

**Abstract:** Many researches have been conducted in various part of the world on women representation in the national legislatures of different countries. They indicate that democracy, type of electoral system, quotas and socio-economic factors influence the systems of legislative bodies. In this study we tested the impacts of seven explanatory variables (political rights, civil liberty, election system type, quota, literacy rate, per-capita income and women participation in labor force) on women representation using multiple linear regression analyses. Data collected from the freedom house and economic surveys were divided into two sub-samples such as training and test data set to develop and test the regression models. Using SPSS version 12 we first constructed the multiple linear regression model and then tested its validity. The results confirm that the model is perfectly fit for explaining variation in women representation. Per-capita income is the only significant predictor which explains 99.5% variation in women representation.

**Key words:** Multiple linear regression analyses, Pakistan, women representation

### INTRODUCTION

The scholars around the world are of the view that despite constituting half of the world population, the representation of women in the national legislature is only 18% and may be regarded as marginal (Krook, 2010; Bano, 2009; Wolbrecht and David, 2007). Even in advance societies the representation of women has increased only marginally from 9% in 1995 to 16 percent in 2004 which is much lower than the critical mass of 30 percent envisaged at Beijing (Paxton *et al.*, 2010; Devlin and Robert, 2008). However, recently it reached to 45 percent especially in Nordic countries, 19.9% in OSCE member countries (excluding Nordic countries) and 22.2% in American societies in 2010 (IPU, 2010). The representation of women in most of the developing countries is much below than those of advanced countries and is about 10.1% in Arabs, 13.2% in Pacific, 18.4% in Sub Saharan Africa and 18.7% in Asian countries (IPU, 2010).

Eminent scholars explained such variations by a variety of factors including institutional factors such as nature of government, electoral system, party system and quotas (Paxton *et al.*, 2010; Stockemer, 2008; Francechet and Jennifer, 2008), socio-economic factors like class-distinction, illiteracy, per capita income, poverty (Jabeen and Muhammad Zafar, 2009) and cultural factors such as customs, traditions and religion (Bano, 2009). The factors responsible for women's overall under-representation is largely common across nations while the degree of under-

representation varies from nation to nation (Pippa and Ronald, 2001). Therefore, common measures and techniques cannot bring similar outcome for all nations, rather customized strategies are mandatory to overcome the problems of women's under representation (Halder, 2004). Although considerable work has been done on women's representation around the world as an individual country case studies (Jabeen and Muhammad Zafar, 2009; Halder, 2004; Lovenduski and Pippa, 2003) or cross countries analysis (Paxton, 1997) but none of the researcher has developed and tested the model with out-of-sample data particularly in Pakistani perspective. This might be one of the main reasons that why different studies with different samples have reached to different results. This paper will fill this gap by developing and testing a model relating to women representation in Pakistan with out-of sample data.

The main objective of this research is to:

- Evaluate trends of women representation in legislature of Pakistan using multiple regressions
- Model relationship between independent and dependent variables and to predict variation in women representation
- test the validity of the developed model

### METHODOLOGY

**Hypotheses:** Among the various factors responsible for variation in women representation, the four major

variables such as democracy, electoral system, quota and socio-economic factors have been included in this study.

It is well documented that democratic form of government provides ideal environments for socially legitimizing and sustaining gains in women's rights and gender equality (Krook, 2010; Bano, 2009). These studies claim that both high early levels of democracy and overtime growth in democracy may increase women's political representation (Paxton *et al.*, 2010). Early democracies through the provision of an early political rights and civil liberty enabled the women to enter in political arena and gain representation. Freedom of thought and expression not only provides women with immediate experience in the political dimensions, but may offer enduring effects that continue over many years by setting women on to a path of higher growth in representation (Jabeen and Muhammad Zafar, 2009). The process is one of increasing returns over time, where women's early presence creates momentum for further gains in the political sphere (Stockemer, 2008).

H 1: High and early democracies will facilitate more women representation in national legislature

There are three mostly used electoral system- Proportional, semi-proportional and Plurality. Most of the previous researches indicate that women representation is higher under proportional representation (PR) than under semi proportional and Plurality systems (Krook, 2010; Bano, 2009; Jabeen and Muhammad Zafar, 2009; Paxton *et al.*, 2010; Stockemer, 2008; Schwindt-Bayer and Mishler, 2005). In a proportional closed list system voter usually votes for party list candidates than an individual candidates, if a party receive 20% votes it means it has secured 20% seats in the legislature. The proportional system is more beneficial for women as due to differences in socioeconomic status, occupational choice and family responsibilities, in comparison to men, women candidates are likely to have greater difficulties in becoming eligible and aspiring political candidature (Inglehart and Pippa, 2003; Kenworthy and Malami, 1999). In such circumstances women have limited chance to enter in to political arena. Thus, because PR electoral systems increase women's chances of recruitment and elect ability despite lower placement on candidate lists, these electoral systems provide greater opportunity for the election of women.

H 2: Women representation will be higher under a more proportional electoral system than a system that is more plurality based.

It is usually argues that women earn less than men, they are less educated, there are fewer women than men in policy making positions and they are more subject to

physical and sexual violence than their counterparts. For tackling this issue many (but not all) of the countries have started to operate some sort of quota system. A quota is a party or legislative rule to fix certain number of seats for women in national legislature. It is well documented that women representation is higher in countries where provisions of some kind of quota/ reservation of seats are available (Murray, 2010; Paxton *et al.*, 2010; Bano, 2009). Scholars generally find that the stronger the gender electoral quota system the greater is the level of women's percentages in political office (Dahlerup, 1998; Caul, 1999; Matland, 1993; Studlar and McAllister, 1998).

H3: Women representation will be higher in countries with some kind of female quota than the countries without it.

Illiteracy, limited access to education, poverty, economic dependency and mobility constraints are the various socio-economic factors that hinder women from the political process of a country (UNDP, 2000). Traditionally the religious and cultural environment of the country restricts women to domestic than to political matters (Azizah, 2002). Predominantly male society doesn't allow women to work outside rather retained them in the boundaries of house to work for the wellbeing of family members. Consequently, women had little opportunities to develop their skill in public domain (Ariffin, 1999). Naturally it is more comfortable to look the household and family affairs while working as house wife and allow the male folk to take the responsibility of the outside world which is relatively more challenging and involves a lot of risk at almost all stages. Thus it is a matter of mutual consensus rather a forced option by men. Of course, not all the females are strong enough to face the risks and challenges of the outside world. Women working outside their houses either for job or social or political work have accepted this challenge in addition to their irreplaceable responsibilities of child bearing and their brought up as well as entertainment of their husbands which is not so easy for majority of the females. Such a model brings mutual satisfaction to the couple but at the cost of low per capita income. Consequently that is one of the major factors for their poor representation in politics and legislative assemblies.

H4: Limited socio-economic and cultural restrictions will ensure restricted women's representation

#### **Data and model:**

**Data:** This article analyzed the growth of women representation at the federal legislature of Pakistan. We evaluated the growth rate of women representation in lower house (National Assembly) of the country for 10 years. The data are drawn from the Inter-Parliamentary

Union (IPU, 2010) and Inter-Parliamentary Union's PARLINE data base supplemented additionally by the work of (Bano, 2009; UNDP, 2007-08) and the government websites. The study was conducted at School of Communication Universiti Sains Malaysia during post doctoral research fellowship in 2010-11.

To assess the impacts of factors on the percentage of women representation, four predictors such as democracy, electoral system, quota/seat reservation and socio-economic factors were included. The dependent variable of the study was the percentage of women representation in the National Assembly and the yearly data was obtained from IPU (2010) and the government websites. For the first independent variable - democracy - we retrieved yearly data from the Freedom (2010). Freedom House provides detail report of annual rating of the entire world from 2002 to date. It is considered authentic and widely used by the researchers of both developed and developing societies (Paxton *et al.*, 2010; Stockemer, 2008). It further breaks rating on the basis of political rights and civil liberty (Freedom, 2010) which helps the researchers to see the impacts of democracy on women representation.

For an electoral system type we used three categories which exist in Europe such as (proportional, semi-proportional and plurality). For regression analysis we created dummy variables and coded one for plurality and zero for proportional and semi-proportional. Data were obtained from the database of Electoral System Design provided by the International Institute for Democracy and Electoral Assistance (IIDERA, 2010).

For the explanatory variable Quota/seat reservation, the overtime reservation of seats of the Pakistani government were obtained from the study of (Bano, 2009) supplemented by the websites of election commission. We used two indicators such as literacy rate and labour force participation to measure the third predictor socio-economic impacts. Data about all these factors were gathered from Human Development Index, established by the United Nations Development Programme supplemented additionally by data drawn from economic surveys available at the website of the Ministry of Finance of Pakistan.

**Modeling:** We used parsimonious models for our regression modeling. A parsimonious model is widely used prediction technique that fit the data adequately without using unnecessary parameters and produces more accurate results (Sonmez, 2004). Prior research indicates that model with a few significant predictors forecast the situation more accurately than the models with insignificant predictors (Abu Bakar and Tahri, 2009; Sonmez, 2004). For finding parsimonious model backward elimination method was used. All the independent variables were first included in the initial

regression model and then irrelevant variables were eliminated one by one. Regression statistics p-value and coefficient of determination ( $R^2$ ) were used as parameters for the elimination of variables. The p-value shows the significance of the variables while  $R^2$  shows variability explained by the model. Ten years data (1999, 2002-2010) on women representation from Pakistan were separated in two sub-samples - the training (6 years) from 1999 and 2002 to 2006 and testing (4 years) from 2007 to 2010. For getting clear picture of overtime variation in women representation we used the data for the year 1999. The non-existence of National Assembly due to declaration of Martial Law in the country in October 1999 was the cause for the exclusion of the data for the year 2000 and 2001. The training data (60%) was used for model building and the testing (40%) data was used for model validation which is higher than the recommended 30% of the sample size for model validation (Kutner *et al.*, 2004).

In order to find parsimonious regression model we initially included the entire variables (political rights; civil liberty; election system type, quota, literacy rate; women participation in labour force; Gross Domestic Product (GDP) and per-capita income in our regression model. Two of them (status of country, election process type) were dummy variables. By using backward methodology we excluded all the irrelevant factors one by one and finally got workable first regression model with four predictors such as quota, literacy rate, labour force participation and per-capita income, which is presented in Eq. (1):

$$PWR = \beta_0 + \beta_{1QUO+} \beta_{2PCI+} \beta_{3LFP+} \beta_{4LR+} \epsilon \quad (1)$$

where PWR is the percentage of women representation; QUO is the quota reserved for women; LR is the literacy rate of women; LFP is the participation rate of women in labour force; PCI is the per-capita income of a country;  $\beta_0, \dots, \beta_4$  are the regression coefficients and  $\epsilon$  is the error term. Regression statistics for model 1 (RM 1) are given in Table 1 with  $R^2$  and the variable corresponding to the coefficient with high p-value. The p-value of 0.996 is for the literacy rate, which indicates the irrelevancy of the variable with the dependent variable hence excluded to get parsimonious model.

The second regression model consisting of three variable such as quota; labour force participation and per-capita income was similar to the first model but do not include literacy rate. The p-values in Table 2 indicate that the contribution of quota to the percentage of women representation is more significant than the contribution of labour force participation and per-capita income. The  $R^2$  value for model 2 remained the same irrespective of the elimination of literacy rate. In the second model  $R^2$ -value

Table 1: Regression models

Models	Independent variables	R <sup>2</sup>	Variables corresponding to the coefficient with the heist p-value	p-value of the coefficient
RMI	QUO, PCI, LFP, LR	1.00	LR	0.996

Table 2: p-values for regression model (RM2)

Independent variables	p-values of the coefficient
(QUO) Quota	0.000
(PCI) Per-capita Income	0.016
(LFP) Labour Force Participation	0.030

Table 3: Regression models (testing)

Models independent variables	R <sup>2</sup>	Variables corresponding to the coefficient with the heist p-value	p-value of the coefficient
RMI QUO, PCI	0.998	QUO	0.427

Table 4: Multicollinearity statistics

		Coefficients <sup>a</sup>					Collinearity stistics	
		Unstandardized coefficients		Standardized coefficients				
Modle		B	S.E	β	T	Sig.	Tolerance	VIF
1	(Constant)	17.029	0.296		63.246	0.010		
	QUOTA2	1.931E-03	0.002	0.068	1.262	0.427	0.604	1.656
2	per capita income	4.949E-03	0.000	1.041	19.320	0.033	0.604	1.656
	(Constant)	17.252	0.231		74.790	0.000		
	per capital income	4.745E-03	0.000	0.998	20.934	0.002	1.000	1.000

a. Dependent Variable: POWINA

of 1.00 indicates that 100% variation in women representation is explained by the three independent factors. There may be certain significant nonlinear or interaction relations between dependent and independent variables that had not been included in the model. For testing the relationship between women representation and its determinants, the proposed multiple regression model is given in Eq. (2).

$$PWR = \beta_0 + \beta_{1QUO} + \beta_{2PCI} + \beta_{3LFP} + \epsilon \quad (2)$$

where PWR is the percentage of women representation; QUO is the quota reserved for women; LFP is the participation rate of women in labour force; PCI is the per-capita income of a country;  $\beta_0, \dots, \beta_3$  are the regression coefficients and  $\epsilon$  is the error term.

**Model validation:** The final stage in model building is the validation of proposed model. Model selected for validation is given in Eq. (3).

$$\hat{Y} = 3.027 + 1.072 + 3.738 - 8.73 \quad (3)$$

$$(R^2 = 1.00, p = 0.000)$$

**Following are the well documented methods for the validation of proposed regression model:** Collection of new data to check the model and its predictive ability. Comparison of results with theoretical expectation, earlier empirical results or simulation results. Use of a hold-out sample to check the model and its predictability (Ismail *et al.*, 2010)

## RESULTS AND DISCUSSION

We used the first method to check the validity of our proposed regression model. Testing data was used to test the fitness of the proposed model. Multiple Linear Regression analysis was used for the validation of proposed mode. We initially included all the three predictors of the proposed model in testing stage but after the exclusion of predictor per-capita income we got the first regression model. The results of the first regression model are given in Table 3. However, for finding parsimonious model we eliminated the insignificant predictors in model 2 through backward methodology. The p-value and coefficient of determination R<sup>2</sup> were used as parameters for the elimination of insignificant variable. After deleting insignificant variable we got the following final regression model RM2:

$$\hat{Y} = 17.252 + 4.745 (R^2 = 0.995, p = 0.002) \quad (4)$$

The results indicate that the model is good enough to predict women representation in national legislature. Total variation explained by the model is about 99.5% which is almost equal to the R<sup>2</sup> = 1.00 value of the training data. Normality of residual is one of the important assumptions of multiple regressions. We tested the normality of the residual of our data and found it normally distributed (Fig. 1).

In addition to the assumption of normality of the residual multiple regressions has the assumption of non-multicollinearity among the independent variables. The high correlation of two or more independent variables

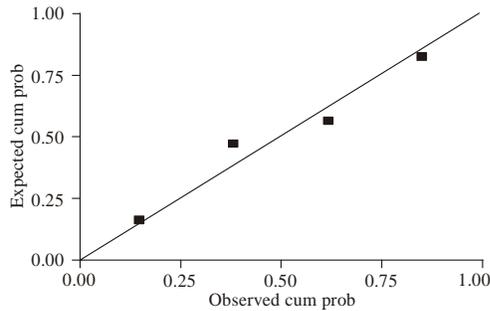


Fig. 1: Normal P-P plot of regression standardized residual

makes it difficult to separate their effect on the dependent variable. Since we had only one independent variable, therefore, there is no multicollinearity problem. Variance Inflation Factors (VIF) higher than 10 is considered problematic which is ok in our case (Table 4).

### DISCUSSION

A galaxy of researchers has unearthed striking results from similar studies and reached similar as well as different findings. For instance the political factors (election system type, quota system and women political rights) socioeconomic factors (education, labor force participation, per capita income) and cultural (religion) factors have been reported over and over as the significant predictors of women representation (Paxton *et al.*, 2010; Stockemer, 2008; Francechet and Jennifer, 2008). The proportional system is more conducive to increasing the role of women in parliamentary affairs (Krook, 2010). Similarly, where there is predefined quota for women, their representation is higher as compared to where there is no predetermined quota for women (Murray, 2010; Dahlerup, 1998; Caul, 1999; Matland, 1993; Studlar and McAllister, 1998). Likewise, if constitution provides well-defined political rights for every group of citizens there are more chances of increase in women representation (Jabeen and Muhammad Zafar, 2009). Similarly highly educated citizens participate more in politics than less educated class (Murray, 2010) and women who find themselves in the formal wage labour force are more likely to enjoy political representation (Kenworthy and Malami, 1999).

In Pakistan, however, the electoral system is based on plurality principle which definitely hinders in increasing women' role but due to quota system introduced during 2002, the women representation has potentially increased (Jabeen and Muhammad Zafar, 2009). We have tested and proved the same in the training-stage of the above given model but lost its explanatory power in testing stage of the model since quota remained unchanged in that period. With regard to political rights, Pakistan stands at the 6<sup>th</sup> position on seven-levels of grading for political rights by

Freedom (2010) therefore it have no impacts what so ever on the growth rate of women political representation. Findings of the study are not supporting the impacts of education on women representation in Pakistan as reported by (Murray, 2010). Similarly the influence of high labor force participation on the growth of women representation found by Inglehart and Pippa (2003) has not been substantiated in the study. Pasha *et al.* (1999) found strong correlation between the level of per capita income and the equalization of economic opportunity between men and women, which is proved by this study.

### CONCLUSION

Given the above literature and analysis, it can be concluded that women representation is dependent on multiple factors, which increase or decrease the role of women in political matters of a country however, their significance and prediction power varies from situation to situation. The reason to fact is that every state provides different social, political, economic and governmental role to different sectors of society. In developing countries like Pakistan the women are still in the backyard and treated as such in almost all aspects of life. If any country is sincere in giving due role to their women in the parliamentary affairs, she must take up the matter on all fronts particularly, proportional electoral system, sufficient quotas, clear cut human rights for women, more opportunities of education & labor force participation and high per capita income of women in a country.

### ACKNOWLEDGMENT

We are thankful to Higher Education Commission (HEC) Pakistan for funding post doctoral study. For their suggestions, assistance, editing and helpful comments, we would like to thanks the anonymous reviewers, editorial board of Maxwell Scientific Organization, as well as colleagues at Gomal University, Dera Ismail Khan, Khyber Pakhtunkhaw, Pakistan.

### REFERENCES

- Abu Bakar, N.M. and I.M. Tahri, 2009. Applying multiple linear regression and neural network to predict bank performance. *Inter. Bus. Res.*, 2: 176-183.
- Ariffin, R., 1999. Feminism in Malaysia: A historical and present perspective of women's struggles in Malaysia. *Women Stud. Inter. Forum*, 22: 417-423.
- Azizah, W., 2002. Women in Politics: Reflections from Malaysia. Retrieved from: [http://www.idea.int/publications/wip/upload/CS\\_Malaysia\\_Azizah.pdf](http://www.idea.int/publications/wip/upload/CS_Malaysia_Azizah.pdf), (Accessed on: May 25, 2010).
- Bano, S., 2009. Women in Parliament in Pakistan: Problems and Potential Solutions. *Women Stud. J.*, 23: 19-35.

- Caul, M., 1999. Women's representation in parliament: The role of political parties. *Party Polit.*, 5: 79-98.
- Dahlerup, D., 1998. Using quotas to increase women's political representation in women in parliament: Beyond the numbers. *Inter-Parliamentary Union*.
- Devlin, C. and E. Robert, 2008. The effect of increased women's representation in parliament: The case of Rwanda. *Parliam. Affairs*, 61: 237-254.
- Francechet and Jennifer, 2008. Gender quotas and women representation. *Polit. Gender*, 4: 393-425.
- Freedom, H., 2010. Freedom in the World. New York: Freedom House. Retrieved from: <http://www.freedomhouse.org/template.cfm?page=15>, (Accessed on: May 26, 2010).
- Halder, N., 2004. Female representation in parliament: A case study from Bangladesh. *New Zealand J. Asi. Stud.*, 6: 27-63.
- Inglehart, R. and N. Pippa, 2003. Gender Equality and Cultural Change. John F Kennedy School of Government, Harvard University, UK.
- Inter Parliamentary Union (IPU), 2010. Women in National Parliaments: Situation as of Inter-Parliamentary Union. Retrieved from: <http://www.ipu.org/wmn-e/classif.htm>, (Accessed on: May 24, 2010).
- International Institute for Democracy and Electoral Assistance, (IIDERA), 2010. Electoral Processes. Retrieved from: <http://www.idea/elections/index.cfm>.
- Ismail, Z., A. Yahya and A. Sabri, 2010. Forecasting gold prices using multiple linear regression model. *Am. J. Appl. Sci.*, 6: 1509-1514.
- Jabeen, N. and J. Muhammad Zafar, 2009. Gender and local governance in Pakistan: representation vs. participation. *Inter. NGO J.*, 4: 264-276.
- Kenworthy, L. and M. Malami, 1999. Gender inequality in political representation: A worldwide comparative analysis. *Social Forces*, 78: 235-269.
- Krook, M.L., 2010. Women's representation in parliament: A qualitative comparative analysis. *Polit. Stud.*, 58: 886-908.
- Kutner, M.H., C.J. Nachtsheim and J. Neter, 2004. *Applied Linear Regression Models*. 4th Edn., McGraw Hill/Irwin, Chicago, IL.
- Lovenduski, J. and N. Pippa, 2003. Westminster women: The politics of presence. *Polit. Stud.*, 51: 84-102.
- Matland, R., 1993. Institutional variables affecting female representation in national legislatures: The case of Norway. *J. Polit.*, 55: 737-755.
- Murray, R., 2010. Second among unequal's? A study of whether France's "Quota women" are up to the job. *Polit. Gender*, 6: 93-118.
- Paxton, P., 1997. Women in National Legislature: A Cross National Analysis. *Social Sci. Res.*, 26: 442-464
- Paxton, P., M. Melanie and M.A. Painter, 2010. Growth in women's political representation: A longitudinal exploration of democracy, electoral system and gender quotas. *Euro. J. Polit. Res.*, 49: 25-52.
- Schwindt-Bayer, L.A. and W. Mishler, 2005. An integrated model of women's representation. *J. Polit.*, 67: 407-428.
- Sonmez, R., 2004. Conceptual cost estimation of building projects with regression analysis and neural networks. *Can. J. Civil Engine.*, 31: 677-683.
- Stockemer, D., 2008. Women's representation: A comparison between Europe and the Americas. *Polit.*, 28: 65-73.
- Studlar, D. and I. McAllister, 1998. Candidate gender and voting in the 1997 British general election: Did labour quotas matter?. *J. Legislative Stud.*, 4: 72-91.
- UNDP., 2000. *Human Development Report*, Oxford University Press, New York.
- UNDP., Report, 2007-08. *Fighting Climate Change: Human Solidarity in a Divided World*. Retrieved from: <http://hdr.undp.org/en/reports/global/hdr2007-2008/>.
- Wolbrecht, C. and E.C. David, 2007. Leading by example: Female members of parliament as political role models. *Am. J. Polit. Sci.*, 51: 921-941.