# Female Literacy and its impact on Demographic & Economic Factors in Maharashtra: Multivariate Approach and Factor Analysis

**Anjum Ara M. K. Ahmad**

Dept of Mathematics & Statistics, Rizvi College of Arts, Science & Commerce, Bandra (W), Mumbai 400050, Maharashtra, India

## Abstract

The objective of the study is to analyse the impact of female literacy rate (FLR) on mean number of births (MCB), infant mortality rate (IMR) and female work participation rate (FWPR) using controls for other demographic and economic variables in the state of Maharashtra. Along with female literacy the effect of various other factors like urban population, population below poverty line, girls marrying below legal age of marriage, per capita gross district domestic product, and religion is also estimated in the present study. To understand the effect of economic status and medical facilities and its utilisation on the dependent variables I developed the two indices, the economic wealth index and the medical index. I have used the district level amenities data and medical facilities data given by Census and other agencies to develop the economic wealth and medical indices respectively. For analysing the data I have used the statistical package SPSS (Version 19). The principal component method is used to develop the economic wealth index and the composite weighted average method is used to develop the medical index. The effect of these indices along with female literacy and other factors on the dependent variables is also studied. Instead of modelling the fertility, infant mortality and female work participation separately, the dependent variables are modelled jointly using various multivariate multiple regression models. The results of Maharashtra State pointed that all the three dependent variables are significantly correlated with each other. It has been observed that the factors which are having a significant effect on the dependent variables are female literacy rate, urbanisation, poverty level, per capita gross district domestic product, religion and medical facilities. Urbanisation is found to have the most significant effect on female work participation. Female literacy is having a strong effect on mean number of births and infant mortality rate in most of the models.

Keywords: Female Literacy, Demography, Economic Factors, Multivariate Approach, Factor Analysis, Maharashtra, India

## 1.0 Introduction

Literacy is the process that liberates the mind. It is liberation from all forms of darkness and ignorance. The effects of female schooling are particularly important for policy-makers because promoting girls' education is a central development strategy (Schultz, 2002). Female literacy is promoted as a policy to increase household productivity by reducing fertility and improving child health, as well as a strategy to build the labour force (Summers, 1994).

Literacy programs primarily aim at improving the quality of human capital, which is very vital for the development of any nation. The confidence and behaviour pattern of the neo-literates change tremendously when they are able to read & write. Freire (1987) a famous Brazilian educator observed that literacy enabled adults to read the world by

reading the word. Literacy can affect people's lives, both women's and men, through several channels. It increases access to knowledge, information and new ideas. It enhances overall efficiency, opportunities in the labour market and social status (Jejeebhoy, 1996).

**1.1 Important Definitions:** Some important definitions used in the study are stated below.

**1.1.1 Literacy Rate:** Literacy is one of the important social characteristics on which information is obtained of every individual in the Census. For the purposes of census a person aged seven and above, who can both read and write with understanding in any language, is treated as literate.

**1.1.2 Mean Number of Births (MNB):** Mean Number of Births is the mean number of children ever born to ever married women who has completed their reproductive cycle. Children ever born to ever married women are given by age group in the Census Data. In order to calculate the mean number of births we divide the "number of Children Ever Born" by "the number of Ever Married Women in the age group (45-49)".

**1.1.3 Infant Mortality Rate (IMR):** The deaths under one year of age are called infant deaths. The conventional measure of infant mortality rate (IMR) is defined as the number of infant deaths occurring in the same community within a specified calendar year per 1000 live birth in the same community during the same calendar year.

**1.1.4 Work Participation Rate (WPR):** Those workers who had worked for the major part of the reference period (i.e. 6 months or more) are termed as Main Workers. Those workers who had not worked for the major part of the reference period (i.e. less than 6 months) are termed as Marginal Workers. Work participation rate is defined as the percentage of total workers (main and marginal) to total population. Work participation rate can be further classified by gender into male work participation rate (MWPR) and female work participation rate (FWPR).

**2. Rationale of the study**
1. Literacy is generally recognized as an important input for social upliftment, especially in areas of population control, health, hygiene environmental degradation control, empowerment of women and weaker sections of the society.
2. The provisional results of Census (2001) corroborate observations that female education strongly influences fertility. This is however a broad trend observed from the provisional result and need further investigation for confirmation.
3. It is generally believed that children born to literate mothers are less likely to experience higher mortality rates than others. A study to check such interrelationship in the selected state is needed.
4. Furthermore there are hardly any study based on Census (2001) data, studying the impact of female literacy trends on the fertility, infant mortality and female work participation in the State of Maharashtra. Such studies can be used for pointing out the problem areas so that policies can be framed for the social and economic development of these regions.

## 3  Objectives of the study

The study involves the following aspects:

1.  To compute the correlation among the factors.
2.  The Factor Analysis using Principal Component Analysis (PCA) method is done to study the importance of the factors.
3.  To assess the impact of female literacy rate (FLR) on mean number of births (MNB), infant mortality rate (IMR) and female work participation rate (FWPR) in the State of Maharashtra.
4.  To estimate along with female literacy the effect of various other factors like urban population, religion, population below poverty line, girls marrying below legal age of marriage, per capita gross district domestic product, economic- wealth and medical index on MNB, IMR and FWPR in the State of Maharashtra.

## 4. Research Methodology:

The research methodology used is explained below.

### 4.1 Collection of Data

The present research is based on secondary data collected from Census data. An attempt is made in this research to study the impact of female literacy on select economic and demographic variables such as fertility, infant mortality and female work participation in Maharashtra.

The district level mean number of births (MNB) and female work participation rates (FWPR) are calculated using the Census (2001) data and included as dependent variables in the study. I have used MNB as an estimate to measure fertility in the study.

Reduction in infant mortality rate (IMR) is also one of the goals of National Rural Health Mission (2005-2012). Hence it is included as one of the dependent variable in this study. Pathak et al. (1988) suggested a method to estimate IMR and CMR from the data on children ever born and children surviving by age of the mother. A revised version of the Pathak's methodology has been used for providing the present estimates. The district level estimates of IMR and CMR for the year 2001 is taken from the book "Infant and Child Mortality in India: District Level Estimate" by Rajan et al. (2008). The district wise data of IMR based on Census 2011 is not available for all the selected states and hence could not be incorporated in the study.

Urbanization is another factor which could affect the dependent variables; hence percentage of urban population is included as one of the independent variables. This data can also be easily calculated from the Census data. As data is available on religious heterogeneity (i.e. population belonging to the various religions) in the Census (2001), this data is included in the study. Percentage of Population of Hindus and Muslims are calculated using Census data and also included as one of the independent variables.

The percentage of population below the poverty line is also considered as one the factors in the study. This variable is available as part of the National Human Development Report (2001). District level values of percentage of people lying below poverty line are given by below Poverty Line Census for 9th plan period by India (1997-2002) (Ministry of Rural Development and Monitoring Division). District-wise Per Capita Gross Domestic Product given by the Planning Commission of India (2001) is another important factor included in this study.

As data is available in the Census on percentage of population owning different assets like radio, transistor, television, telephone, bicycle, scooter, motor cycle, moped,

car-jeep-van, this data is used in the study for assessing the economic health of the area. According to National Human Development Report (2001) the available Census data permits analysis of two aspects of quality of housing and shelter namely, living space or the number of rooms available to a household and the quality of construction of the residence i.e. whether a household resides in a pucca or a kutcha construction. The Census also presents data on quality of houses based on the material used for construction of walls and roof separately. There are also large inter-State variations in the availability of electricity, access to toilet facilities and safe drinking water at the household level, both in urban and in rural areas. Also the Indian Government has for long been using electronic and other mass media to promote family planning, child health and women empowerment hence the ownership of television and radio in a large segment of the population can possibly have a salutary impact on the MNB, IMR and FWPR. An attempt is made to study the effect of these assets on the dependent variables by constructing the Economic Wealth Index.

Information on the availability and utilisation of health facilities are available at the district level. These data are given in Census reports, District Level Household and Facility Survey (DLHS) and Reproductive and Child Health (RCH) Projects conducted by Registrar General of India and other bodies like International Institute for Population Sciences (IIPS, Mumbai). To study the effect of medical facilities on the dependent variables the medical index is constructed using these data.

### 4.2 Economic Wealth Index

The Economic Well-Being Index (EWI) using Principal Component Method (PCA) is constructed in this study based on the amenities data, condition of houses and presence of various basic facilities like drinking water, electricity and toilet given by Census. This index is used as control variable in the regression analysis.

### 4.2.1  Principal Components Analysis

PCA is a multivariate statistical technique used to reduce the number of variables in a data set into a smaller number of 'dimensions'. In mathematical terms, from an initial set of $n$ correlated variables, PCA creates uncorrelated indices or components, where each component is a linear weighted combination of the initial variables. For example, from a set of variables $X_1$ through to $X_n$,

$$PC_1 = a_{11}X_1 + a_{12}X_2 + \ldots\ldots\ldots a_{1n}X_n$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$PC_m = a_{m1}X_1 + a_{m2}X_2 + \ldots\ldots a_{mn}X_n$$

where $a_{mn}$ represents the weight for the $m^{th}$ principal component and the $n^{th}$ variable, i.e. in brief if there are n correlated variables $X_1 \ldots.. X_n$, each principal component (PC) is the sum of each variable multiplied by its weight (the weight for each variable is different in each principal component).

The weights for each principal component are given by the eigenvectors of the correlation matrix, or if the original data are standardized, the co-variance matrix. The variance ($\lambda$) for each principal component is given by the Eigen value of the corresponding eigenvector. The components are ordered so that the first component (PC$_1$)

explains the largest possible amount of variation in the original data. The second component ($PC_2$) is completely uncorrelated with the first component, and explains additional but less variation than the first component. Subsequent components are uncorrelated with previous components; therefore, each component captures an additional dimension in the data, while explaining smaller and smaller proportions of the variation of the original variables. The higher the degree of correlation among the original variables in the data, the fewer components required to capture common information. Before applying the PCA following tests are necessary:

**a) Kaiser-Meyer-Olkin Measure of Sampling Adequacy:** This measure varies between 0 and 1, and values closer to 1 are better. A value of 0.5 is a suggested minimum.

**b) Bartlett's Test of Sphericity**: This tests the null hypothesis stating that the correlation matrix is an identity matrix. An identity matrix is one in which all of the diagonal elements are 1 and all off diagonal elements are 0. This null hypothesis should be rejected.

Taken together, these tests provide a minimum standard which should be passed before principal components analysis (or factor analysis) should be conducted.

**4.2.2 Composite Index:** The indexing of districts using a set of variables has multiple uses. It helps in assessing current level of development, monitor the trends and identify target areas and groups requiring special attention not only by the country but in each state and district as well. The composite index is computed after standardizing the variables and by fixing the upper and lower limit.

**4.2.2.1 Standardization of Indicators :** The indicators may be grouped into positive and negative indicators. For example coverage of complete immunization is a positive indicator while proportion of births of order three and above is a negative indicator. In computing composite indices I have to make them uniform and therefore there is a need of standardizing each indicator.

For positive indicators, composite index is usually computed as
100* (Vi – Vmin) / (Vmax-Vmin).

Here Vi is Actual Value in the series, Vmin is lowest value observed in the series, Vmax is highest value observed in the series. This simply states that the district with a lowest value will get a score of 0 and district with a highest value will get a score of 100.

In case variables affect negatively, composite index is computed as
100* (Vmax-Vi) /(Vmax-Vmin).

**4.2.3 Selection of Variables:** The data of amenities which are available district wise in Census (2001) are used for the construction of the Economic Wealth Index. Percentages of households possessing these amenities are calculated from Census (2001) data. To estimate the economic wealth index, principal component analysis based on the first principal component is used.

Following Nineteen variables have been selected for developing the index from the amenities district wise data of the four selected states.

1) Percentage of households having Good condition houses (Gd_con_houses).
2) Percentage of households having Permanent houses (Per_perm).
3) Percentage of households having Separate kitchen (Per_sepkit).
4) Percentage of households having LPG connection (Per_LPG).
5) Percentage of households having Electricity (Per_elec).
6) Percentage of households having Water within Premises (Per_water).
7) Percentage of households having Bathrooms (Per_of_bath).
8) Percentage of households having Drainage (Per_of_drain).
9) Percentage of households having Toilets (Per_toilet).
10) Percentage of households having Cement or Mosaic Flooring (Per_cemmosaicfloor).
11) Percentage of households having Concrete Roof (Per_concreteroof).
12) Percentage of households having Brick, Stone and Concrete Walls (Per_brstconcretewall).
13) Percentage of households having Bank accounts (Per_bank).
14) Percentage of households having Radio (Per_radio).
15) Percentage of households having Television (Per_TV).
16) Percentage of households having Telephone (Per_telephone).
17) Percentage of households having Bicycles (Per_bicycles).
18) Percentage of households having Scooter (Per_scootermotor).
19) Percentage of households having Car or Jeep (Per_carjeep).

By definition the first principal component variable across households or individuals has a mean of zero and a variance of X, which corresponds to the largest Eigenvalue of the correlation matrix of X. The first principal component yields a wealth index that assigns a larger weight to assets that vary the most across households so that an asset found in all households is given a weight of zero (McKenzie, 2005). The first principal component can take positive as well as negative values.

The SPSS factor analysis procedure is used in this study. This procedure first standardizes the indicator variables (calculating Z scores); then the factor coefficient scores (factor loadings) are calculated; and finally, for each district, the indicator values are multiplied by the loadings and summed to produce the Economic Wealth Index (EWI) value for each district. In this process, only the first factor produced is used to represent the index. The resulting sum is itself a standardized score with a mean of zero and a standard deviation of one.

**4.2.4 Composite Economic Wealth Index (CEWI)**

Since the EWI index takes positive as well as negative values, the Composite EWI index is computed using the formula

$$\text{Composite EWI} = \frac{(\text{EWI Value} - \text{Minimum EWI Value}) * 100}{(\text{Max EWI Value} - \text{Min EWI Value})}$$

**4.3 Medical Index (MI):** In order to construct this index I have included information on infrastructure and utilisation of health facilities available at the district level from various Surveys like Census, Reproductive and Child Health (RCH) conducted by Registrar General of India and other bodies. Weighted Average Method and Composite Index Method are used to construct the Medical Index and the Composite Medical Index.

**4.3.1 Selection of Variables:** The selection of variables is extremely important in

any such experience. Census 2001 gives us data of number of Sub Centres (SC), Public Health Centres (PHC) and Community Health Centres or Rural Hospitals (CHC). These data are available district wise, with the limitation that these data are available mainly for rural areas. Medical Statistics for Mumbai and Mumbai Suburban district is taken from Mumbai Human Development Report (2009) and Combine Mumbai and Mumbai Suburban data from Health and Healthcare in Maharashtra, a Status Report by Duggal et al.(2005). Also there are data available at district level by the District Level Household (DLHS) Surveys and Reproductive and Child Health (RCH), for the various facilities present at public health centres, percentage of immunisation and percentage of women taking anti natal care. All these data are used to assess the medical facility at the district level by constructing the medical index. In order to maintain homogeneity of data, entire data source centre around the year 2001.

**4.3.2 Development of Medical Index:** This index is created in the following steps:
a) Medical Infrastructure Index: This index is actually the number of beds per lakh population. It is calculated using the weighted average formula. In this method weight-age is given to the centres according to number of beds in it. Therefore sub centre (SC), public health centre (PHC) with lesser beds is given lesser weight-age as compared to community health centre (CHC) in calculating the index. Generally SC is a single bed centre, PHC has 6 beds and CHC (or Rural Hospitals) has about 30 beds. Hence the formula used for calculating Medical Infrastructure Index is
(SC*1+PHC*6+CHC*30)/P *100000
where P stands for total population of the area.

b) Anti Natal Care (ANC) Index: To create this index I have taken the average of three factors (given in percentages) V04, V05 and V06 from the District Level Key Indicators given by District Level Household Survey-2 and Reproductive and Child Health (DLHS2-RCH) data where V04 is women received at least one TT Injection, V05 is women received hundred or more Iron and Folic Acid (IFA) tablets and syrups and V06 is women with at least three ANC Visits. Then the Composite Anti Natal Care Index is calculated using the following formula
Comp ANC Index = (ANC Index Value– Minimum ANC Index Value) * 100
                               (Max ANC Index Value- Min ANC Index Value)

c) Immunisation Index: To calculate this index I have taken the average of three factors (given in percentages) V09, V10 and V11 from the District Level Household Survey and Reproductive and Child Health (DLHS2-RCH) data where V09 is infants given three Injections of DPT, V10 is infants given Injection of Measles, V11 is infants given Full Immunisation (ie.BCG + 3 doses of Polio + 3 injections of DPT + Measles). Then Composite Immunisation Index is calculated using the following formula
Comp Imm Index =   (Imm Index Value– Minimum Imm Index Value) * 100
                               (Max Imm. Index Value- Min Imm Index Value)

d) Public Health Centre (PHC) facility index: In Public Health Centre (PHC) facility index, the average of the three facilities i.e. staff, infrastructure and supply is taken from the DLHS2-RCH data and then the Composite PHC index is calculated using the formula

Composite PHC Index = (PHC Index Value– Minimum PHC Index Value) * 100
$\qquad\qquad\qquad$ (Max PHC Index Value- Min PHC Index Value)

e) Medical Utilisation Index (MUI): This index is constructed by taking the average of Immunisation Index and Anti-Natal Care Index. Further the Composite Medical Utilisation Index (CMUI) is calculated by using composite index formula.

f) Medical Index (MI): The Medical index is the average of four composite indices i.e. Composite medical infrastructure index, Composite PHC index, Composite Immunisation index and Composite Anti-Natal Care index.

### 4.3.3  Composite Medical Index (CMI): To compute CMI the composite index of the Medical Index (MI) is calculated using the formula

Composite Medical Index   =  (MI Value– Minimum MI Value) * 100
$\qquad\qquad\qquad\qquad\qquad$ (Max MI Value- Min MI Value)

### 4.4 Analysis of Data

$\qquad$ To find the significant factors affecting the dependent variables the various multivariate multiple regression analysis models are applied in the study.
The dependent and independent variables used in the study are as follows.
Dependent Variable
$\qquad$ a) Mean Number of Births (MNB)
$\qquad$ b) Infant Mortality Rate (IMR)
$\qquad$ c) Female Work Participation Rate (FWPR)
Independent Variables
$\qquad$ 1) Female literacy rate (FLR)
$\qquad$ 2) Male literacy rate (MLR)
$\qquad$ 3) Percentage of Urban Population in Total Population (PURBAN)
$\qquad$ 4) Percentage of Girls Marrying Below the Legal Age of Marriage (PGMBLA)
$\qquad$ 5)  Population Below Poverty Line (BPL)
$\qquad$ 6)  Per Capita Gross District Domestic Product at (1999-2000) price (GDDP)
$\qquad$ 7)  Composite Economic Wealth Index (CEWI)
$\qquad$ 8)  Percentage of Hindus in Total Population (HINDU)
$\qquad$ 9)  Percentage of Muslim in Total Population (MUSLIM)
$\qquad$ 10) Composite Medical Index (CMI)

### 4.4.1 Multivariate Multiple Linear Regression Model

$\qquad$ The multivariate multiple linear regression model is used to model the relationship between m responses $Y_1, Y_2, \ldots\ldots\ldots, Y_m$ and a single set of predictor variables, $z_1, z_2 \ldots\ldots\ldots\ldots z_r$. Each response is assumed to follow its own regression model so that

$$Y_1 = \beta_{01} + \beta_{11}z_1 + \ldots\ldots\ldots\ldots + \beta_{r1}z_r + e_1$$
$$Y_2 = \beta_{02} + \beta_{12}z_1 + \ldots\ldots\ldots\ldots + \beta_{r2}z_r + e_2$$
$$\cdot \qquad \cdot \qquad\qquad \cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot$$
$$\cdot \qquad \cdot \qquad\qquad \cdot \qquad\quad \cdot \qquad\quad \cdot \qquad\quad \cdot$$
$$Y_m = \beta_{0m} + \beta_{1m}z_1 + \ldots\ldots\ldots\ldots + \beta_{r\,m}z_r + e_m$$

The error term $e' = [e_1, e_2 \ldots\ldots\ldots\ldots e_m]$ has $E(e) = 0$ and $Var(e) = \sum$. Thus, the error terms associated with different responses may be correlated.

To establish notation conforming to the classical linear regression model, let $[Z_{j0}, Z_{j1}, Z_{j2} \ldots\ldots\ldots Z_{jr}]$ denote the values of the predictor variables for the $j^{th}$ trial (j=1,2….n). Let $Yj' = [Y_{j1}, Y_{j2} \ldots\ldots\ldots Y_{jm}]$ be the responses, and let $e_j' = [e_{j1}, e_{j2} \ldots\ldots\ldots e_{jm}]$ be the errors.

In matrix notation, the design matrix is

$$Z_{(nx(r+1))} = \begin{bmatrix} Z_{10} & Z_{11} & Z_{12} & . & Z_{1r} \\ Z_{20} & Z_{21} & Z_{22} & . & Z_{2r} \\ . & . & . & . & . \\ Z_{n0} & Z_{n1} & Z_{n3} & . & Z_{nr} \end{bmatrix}$$

Z is the same as that for the single-response regression model. The other matrix quantities have multivariate counterparts. Set

$$Y_{(nxm)} = \begin{bmatrix} Y_{11} & Y_{12} & . & Y_{1m} \\ Y_{21} & Y_{22} & . & Y_{2m} \\ . & . & . & . \\ Y_{n1} & Y_{n2} & . & Y_{nm} \end{bmatrix} = [Y_{(1)} : Y_{(2)} : \quad : Y_{(m)}]$$

$$\beta_{((r+1)xm)} = \begin{bmatrix} \beta_{01} & \beta_{02} & . & \beta_{0m} \\ \beta_{11} & \beta_{12} & . & \beta_{1m} \\ . & . & . & . \\ \beta_{r1} & \beta_{r2} & . & \beta_{rm} \end{bmatrix} = [\beta_{(1)} : \beta_{(2)} : \quad : \beta_{(m)}]$$

$$e_{(nxm)} = \begin{bmatrix} e_{11} & e_{12} & . & e_{1m} \\ e_{21} & e_{22} & . & e_{2m} \\ . & . & . & . \\ e_{n1} & e_{n2} & . & e_{nm} \end{bmatrix} = [e_{(1)} : e_{(2)} : \quad : e_{(m)}]$$

Hence the multivariate linear model will be:

$$Y_{(n \times m)} = Z_{(n \times (r+1))} \beta_{((r+1) \times m)} + e_{(nxm)}$$

with

$$E(e(i)) = 0 \; ; \quad Cov(e(i), e(k)) = \sigma_{ik}I \qquad i,k = 1,2,\ldots\ldots m$$

The m observations on the $j^{th}$ trial have covariance matrix $\sum = \{\sigma_{ik}\}$, but observations from different trials are uncorrelated. Here $\beta$ and $\sigma_{ik}$ are unknown parameters; the design matrix Z has $j^{th}$ row $[z_{j0}, z_{j1},.......... z_{jr}]$.

**4.4.2 Measures for assessing Statistical Significance:** The four most widely used measures in multivariate multiple regression model for assessing statistical significance between groups on the independent variables are:
a)      Roy's greatest characteristic root(q)
b)      Wilks' lambda(λ)
c)      Pillai's criterion(V)
d)      Hotelling's Trace(U)
In most situations the results/conclusions are same across all four measures but in some unique instances the result will differ between measures.

**5    Analysis of Maharashtra data:**
      The correlation matrix giving the correlation between the dependent and the independent data is calculated. The relationship between FLR and the dependent variables MNB, IMR and FWPR and the other independent variables are studied here. It is necessary to study the correlation between FLR and other independent variables so that I can find the inter-colinearity between the variables. The Factor Analysis using Principal Component Analysis (PCA) method is done to check whether the factors considered are important or not and then finally I have run the Multivariate Regression Analysis. The dependent variables are MNB, IMR and FWPR and the independent factors considered for the study are the three dummy variables (States BIHAR, KERALA and MAHARASHTRA) and the ten continuous variables FLR, MLR, PURBAN, PGMBLA, BPL, CEWI, GDDP, HINDU, MUSLIM and CMI. Jharkhand is taken as reference State among the dummy variables.

**5.1 Correlation Analysis – Maharashtra data**
      In Table 5.1 gives the Correlation Analysis between the dependent and the independent variables in the State of Maharashtra. The correlation between FLR and the other variables is also analysed.

**Table 5.1:  Correlation Analysis of Factors in Maharashtra**

| Correlations Matrix | | | | |
|---|---|---|---|---|
| **Factors** | **FLR** | **MNB** | **IMR** | **FWPR** |
| **MNB** | -.625[*] | 1 | .640[*] | .430[*] |
| **IMR** | -.459[*] | .640[*] | 1 | .396[**] |
| **FWPR** | -.691[*] | .430[*] | .396[**] | 1 |
| **FLR** | 1 | -.625[*] | -.459[*] | -.691[*] |
| **MLR** | .892[*] | -.489[*] | -.509[*] | -.490[*] |
| **PGMBLA** | -.733[*] | .479[*] | 0.103 | 0.242 |
| **PURBAN** | .642[*] | -.513[*] | -0.322 | -.943[*] |
| **BPL** | -0.196 | .432[*] | .710[*] | 0.285 |
| **CEWI** | .748[*] | -.756[*] | -.611[*] | -.856[*] |
| **GDDP** | .718[*] | -.762[*] | -.497[*] | -.797[*] |
| **CMI** | .675[*] | -.737[*] | -.492[*] | -.600[*] |

| | | | | |
|---|---|---|---|---|
| **HINDU** | **-0.263** | **-0.155** | **-0.084** | **.549**[*] |
| **MUSLIM** | **0.113** | **0.088** | **-0.237** | **-.593**[*] |

Source: Derived with the help of SPSS
*1 percent LOS, **5 percent LOS

In the State of MAHARASHTRA all the three dependent variables MNB, IMR and FWPR are significantly correlated with each other. It is observed in Table 5.1 that FLR is significantly and negatively correlated with all the three dependent variable MNB, IMR and FWPR. MNB is significantly correlated with all the given factors except both the Religion HINDU and MUSLIM. IMR is significantly correlated with the factors FLR, MLR, BPL, CEWI, Per Capita GDDP (GDDP) and CMI. FWPR is significantly correlated with the all the given factors except PGMBLA and BPL. The Correlation Matrix points out that most of the independent variables are highly correlated with each other.

**5.2 Factor Analysis - Maharashtra data:** Factor Analysis is done taking the ten economic and demographic factors as variables.
**5.2.1 KMO and Bartlett's Test - Maharashtra data:** The KMO Measure of Sampling Adequacy is 0.638 and the Bartlett's Test of Sphericity is significant. Hence, it is concluded that the data is fit for PCA.

**Table 5.2.1 KMO and Bartlett's Test in Maharashtra data**

| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .638 |
|---|---|---|
| **Bartlett's Test of Sphericity** | **Approx. Chi-Square** | 402.696 |
| | **df** | 45 |
| | **Sig.** | .000 |

Source: Derived with the help of SPSS

**5.2.2 Communalities - Maharashtra data:** From the communalities table it is evident that the maximum variance in the data is due to HINDU (0.921) followed by FLR (0.918) and MUSLIM (0.916).

**Table 5.2.2 Communalities in Maharashtra data**

| Factors | Initial | Extraction |
|---|---|---|
| MLR | 1.000 | .665 |
| FLR | 1.000 | .918 |
| HINDU | 1.000 | .921 |
| MUSLIM | 1.000 | .916 |
| PGMBLA | 1.000 | .915 |
| BPLREG | 1.000 | .799 |
| CEWI | 1.000 | .900 |
| PURBAN | 1.000 | .831 |
| GDDP | 1.000 | .846 |
| CMI | 1.000 | .688 |
| Extraction Method: Principal Component Analysis. | | |

Source: Derived with the help of SPSS

**5.2.3 Total Variance Explained and Scree Plot- Maharashtra data**

Factor analysis is done using Principal component extraction method with varimax rotation. Table 5.2.3 indicate that three principal component factors are extracted whose Eigen values are more than one which explained 83.985 percent of variance in the data.

**Table 5.2.3: Total Variance Explained in Maharashtra data**

| Comp | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | Percent of Variance | Cum percent | Total | Percent of Variance | Cum percent | Total | Percent of Variance | Cum percent |
| 1 | 5.215 | 52.148 | 52.148 | 5.215 | 52.148 | 52.148 | 4.478 | 44.778 | 44.778 |
| 2 | 2.030 | 20.301 | 72.449 | 2.030 | 20.301 | 72.449 | 2.222 | 22.220 | 66.998 |
| 3 | 1.154 | 11.536 | 83.985 | 1.154 | 11.536 | 83.985 | 1.699 | 16.987 | 83.985 |
| 4 | .789 | 7.889 | 91.874 | | | | | | |
| 5 | .434 | 4.344 | 96.219 | | | | | | |
| 6 | .169 | 1.694 | 97.912 | | | | | | |
| 7 | .099 | .993 | 98.905 | | | | | | |
| 8 | .065 | .652 | 99.557 | | | | | | |
| 9 | .027 | .273 | 99.830 | | | | | | |
| 10 | .017 | .170 | 100.000 | | | | | | |
| **Extraction Method: Principal Component Analysis.** | | | | | | | | | |

Source: Derived with the help of SPSS

**Figure 5.2.3: Scree Plot of Maharashtra**
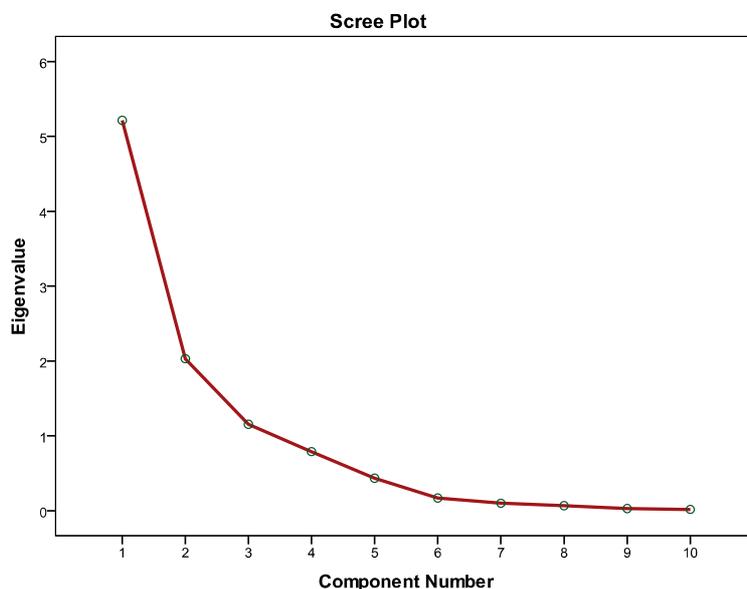


Source: Drawn with the help of SPSS

The Scree Plot (Figure 5.2.3) also verified that the first three principal components are sufficient to explain the variance in the data.

### 5.2.4 Classification of Factors - Maharashtra data

Depending on the loading of the variables in the rotation matrix (Table 5.2.4a), the variables are categorised into three components. MLR and FLR have the highest loading in the matrix on component 1. Hence, it goes to component 1. HINDU and MUSLIM have the highest loading in the matrix on component 2. So, it is placed in component 2. BPL has the highest loading in component 3, that's why it is placed in component 3.

**Table 5.2.4a: Rotated Component Matrix of factors in Maharashtra data**

| Rotated Component Matrix[a] | | | |
|---|---|---|---|
| **Factors** | **Component** | | |
| | **1** | **2** | **3** |
| **MLR** | .808 | .085 | .074 |
| **FLR** | .950 | .119 | .038 |
| **HINDU** | -.171 | -.931 | .158 |
| **MUSLIM** | -.032 | .911 | .292 |
| **PGMBLA** | -.850 | .255 | .357 |
| **BPL** | -.116 | -.042 | -.885 |
| **CEWI** | .745 | .250 | .531 |
| **PURBAN** | .627 | .563 | .347 |
| **GDDP** | .794 | .177 | .429 |
| **CMI** | .760 | .163 | .289 |
| **Rotation Method: Varimax with Kaiser Normalization.** | | | |
| **a. Rotation converged in 6 iterations.** | | | |

Source: Derived with the help of SPSS

The three components are named as Literacy and other factors, Religion and Poverty. It is observed in the factor Analysis that all the variables are important hence all the factors are used in the study.

**Table 5.2.4b: Classification of Factors in Maharashtra**

| Component 1 | Component 2 | Component 3 |
|---|---|---|
| (Literacy and other factors) | (Religion) | (Poverty) |
| ( percent of Variance is 44.778) | ( percent of Variance is 22.22) | ( percent of Variance is16.99) |
| FLR (0.950) | HINDU (-0.93) | BPL (-0.885) |
| PGMBLA (-0.850) | MUSLIM (0.91) | |
| MLR (0.808) | | |
| GDDP (0.794) | | |
| CMI (0.761) | | |
| CEWI ( 0.745) | | |
| PURBAN (0.627) | | |

Source: Derived from Table 5.2.4a

**5.3 Multivariate Analysis- Maharashtra data** To find the significant factors Multivariate multiple regression analysis is run. Instead of MNB and IMR, LOGMNB and LOGIMR are found to satisfy better normality conditions. Hence the dependent variables considered are LOGMNB, LOGIMR and FWPR. To study the effect of female literacy on these dependent variables, multivariate multiple regression model described in the Research Methodology (Point 4.4) is used. In the model the variables are included one after the other i.e. stepwise analysis is done. Variables with high inter-collinearity are dropped from the same model. Also variables which are not significant are dropped from the model. I have tried to find out the best model using significant variables only.

The effect of the independent variables in the models is tested at 5% and 1% level of significance using test statistics like Wilk's λ, Roy's largest root q, Pillai's V and Lawley-Hotelling U. To check whether the model is a good fit or not, appropriate statistics like $R^2$ are used.

The Test of between Subject Effects for the significant models and the linear regression results of the dependent variables are also noted down (using Statistical package SPSS).

When this model is run using SPSS (Version 19) it gives the multivariate analysis result, the individual ANOVA (Test of between Subjects Effects) result of the three dependent variables and the multiple linear regression coefficients with the parameter estimates. Thus, by running multivariate analysis in SPSS, I get all the results at one go and hence I save lot of time. The results of the various Models are summarised in the Tables given below (Tables 5.3 a, 5.3 b and 5.3 c).

**Table 5.3a: Results of Multivariate Analysis in Maharashtra**

| Effect | | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|---|
| | | Value | Value | Value | Value | Value |
| **Intercept** | **Pillai's Trace** | 0.936* | 0.959* | 0.967* | 0.945* | 0.857* |
| | **Wilks' Lambda** | 0.064* | 0.041* | 0.033* | 0.055* | 0.143* |
| | **Hotelling's Trace** | 14.539* | 23.458* | 29.083* | 17.177* | 5.977* |
| | **Roy's Largest Root** | 14.539* | 23.458* | 29.083* | 17.177* | 5.977* |
| **FLR** | **Pillai's Trace** | 0.601* | 0.391* | 0.301** | 0.396* | 0.419* |
| | **Wilks' Lambda** | 0.399* | 0.609* | 0.699** | 0.604* | 0.581* |
| | **Hotelling's Trace** | 1.508* | 0.643* | 0.43** | 0.655* | 0.721* |
| | **Roy's Largest Root** | 1.508* | 0.643* | 0.43** | 0.655* | 0.721* |
| **PURBAN** | **Pillai's Trace** | | 0.894* | 0.798* | 0.842* | 0.841* |
| | **Wilks' Lambda** | | 0.106* | 0.202* | 0.158* | 0.159* |
| | **Hotelling's Trace** | | 8.416* | 3.94* | 5.346* | 5.282* |
| | **Roy's Largest Root** | | 8.416* | 3.94* | 5.346* | 5.282* |
| **BPL** | **Pillai's Trace** | | | 0.483* | 0.503* | 0.358* |
| | **Wilks' Lambda** | | | 0.517* | 0.497* | 0.642* |
| | **Hotelling's Trace** | | | 0.933* | 1.013* | 0.557* |
| | **Roy's Largest Root** | | | 0.933* | 1.013* | 0.557* |
| **GDDP** | **Pillai's Trace** | | | 0.396* | | |
| | **Wilks' Lambda** | | | 0.604* | | |
| | **Hotelling's Trace** | | | 0.655* | | |
| | **Roy's Largest Root** | | | 0.655* | | |
| **HINDU** | **Pillai's Trace** | | | | | 0.26** |
| | **Wilks' Lambda** | | | | | 0.74** |
| | **Hotelling's Trace** | | | | | 0.351** |
| | **Roy's Largest** | | | | | 0.351** |

| | | | | | 0.411* | 0.385* |
|---|---|---|---|---|---|---|
| | Root | | | | | |
| **MUSLIM** | **Pillai's Trace** | | | | **0.411*** | **0.385*** |
| | **Wilks' Lambda** | | | | **0.589*** | **0.615*** |
| | **Hotelling's Trace** | | | | **0.697*** | **0.626*** |
| | **Roy's Largest Root** | | | | **0.697*** | **0.626*** |
| **CMI** | **Pillai's Trace** | | | | | **0.347*** |
| | **Wilks' Lambda** | | | | | **0.653*** |
| | **Hotelling's Trace** | | | | | **0.532*** |
| | **Roy's Largest Root** | | | | | **0.532*** |

\* 1 percent LOS, \*\*5 percent LOS

 Source: Derived with the help of SPSS

Models described in Table 5.3aare as follows

Model 1: Design: Intercept + FLR

Model 2: Design: Intercept + FLR + PURBAN

Model 4: Design: Intercept + FLR + PURBAN + BPL + GDDP

Model 4: Design: Intercept + FLR + PURBAN + BPL+ MUSLIM

Model 5: Design: Intercept + FLR + PURBAN + BPL + HINDU+ MUSLIM + CMI


**Table 5.3 b: Results of Tests of Between-Subjects Effects in Maharashtra**

| **Tests of Between-Subjects Effects of the Models** | | | | | | |
|---|---|---|---|---|---|---|
| | | **Model 1** | **Model 2** | **Model 3** | **Model 4** | **Model 5** |
| **Source ($X_j$)** | **Dept Variables ($Y_i$)** | F | F | F | F | F |
| **Corrected Model** | **LOGMNB[a]** | 21.501* | 11.735** | 17.983** | 14.542* | 18.083* |
| | **LOGIMR[b]** | 9.05* | 4.393* | 15.563** | 12.311* | 12.175* |
| | **FWPR[c]** | 31.013* | 152.566* | 77.818* | 103.066* | 71.104* |
| **Intercept($a_i$)** | **LOGMNB** | 327.371* | 229.825* | 314.306* | 152.354* | 64.311* |
| | **LOGIMR** | 318.689* | 231.171* | 351.944* | 218.731* | 82.254* |
| | **FWPR** | 110.637* | 217.451* | 184.269* | 222.822* | 44.412* |
| **FLR** | **LOGMNB** | 21.501* | 7.698* | 1.854 | 3.491 | 0.546 |
| | **LOGIMR** | 9.05* | 5.077** | 3.43 | 6.838** | 6.986** |
| | **FWPR** | 31.013* | 4.135* | 5.851* | 11.3* | 13.916* |

| | | | * | * | | |
|---|---|---|---|---|---|---|
| **PURBAN** | **LOGMNB** | | **1.594** | **5.862\*\*** | **8.501\*** | **6.971\*\*** |
| | **LOGIMR** | | **0.001** | **5.564\*\*** | **0.423** | **0.768** |
| | **FWPR** | | **143.833\*** | **83.674\*** | **81.875\*** | **84.812\*** |
| **BPL** | **LOGMNB** | | | **3.421** | **13.568\*** | **6.046\*\*** |
| | **LOGIMR** | | | **28.779\*** | **30.914\*** | **15.666\*** |
| | **FWPR** | | | **1.032** | **0.015** | **0.252** |
| **GDDP** | **LOGMNB** | | | **19.776\*** | | |
| | **LOGIMR** | | | **5.109\*\*** | | |
| | **FWPR** | | | **1.929** | | |
| **HINDU** | **LOGMNB** | | | | | **2.579** |
| | **LOGIMR** | | | | | **7.881\*** |
| | **FWPR** | | | | | **1.954** |
| **MUSLIM** | **LOGMNB** | | | | **12.987\*** | **2.832** |
| | **LOGIMR** | | | | **0.072** | **3.759** |
| | **FWPR** | | | | **11.644\*** | **10.952\*\*** |
| **CMI** | **LOGMNB** | | | | | **14.568\*** |
| | **LOGIMR** | | | | | **1.593** |
| | **FWPR** | | | | | **1.228** |
| **a. R² (R²adj)** | **LOGMNB** | **0.387 (0.369)** | **0.416 (0.380)** | **0.699 (0.660)** | **0.652 (0.607)** | **0.789 (0.745)** |
| **b. R² (R²adj)** | **LOGIMR** | **0.210 (0.187)** | **0.210 (0.162)** | **0.668 (0.625)** | **0.614 (0.564)** | **0.716 (0.657)** |
| **c. R² (R²adj)** | **FWPR** | **0.477 (0.462)** | **0.902 (0.896)** | **0.909 (0.898)** | **0.930 (0.921)** | **0.936 0.923)** |

\* 1 percent LOS, \*\*5 percent LOS
Source: Derived with the help of SPSS

The Models are same as defined in Table 5.3a
The Linear Regression Equations for the dependent variables can be defined as
LOGMNB $(Y_1) = a_1 + B_{11}X_1 + B_{12}X_2 + B_{13}X_3 + B_{14}X_4 + B_{15}X_5 + B_{16}X_6 + B_{17}X_7$
LOGIMR $(Y_2) = a_2 + B_{21}X_1 + B_{22}X_2 + B_{23}X_3 + B_{24}X_4 + B_{25}X_5 + B_{26}X_6 + B_{27}X_7$
FWPR $(Y_3) = a_3 + B_{31}X_1 + B_{32}X_2 + B_{33}X_3 + B_{34}X_4 + B_{35}X_5 + B_{36}X_6 + B_{37}X_7$
Where $Y_i$ = the $i^{th}$ dependent variable
$X_j$ = the $j^{th}$ independent variable
$a_i$ = the intercept of the $i^{th}$ dependent variable

Bij = the unstandardised coefficient of the ith dependent variable and j[th] independent variable in the various Models

**Table 5.3c: Multiple Linear Regression B Estimates in Maharashtra**

| Unstandardised B Estimates of Various Models | | | | | | |
|---|---|---|---|---|---|---|
| | | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
| Dept Variable ($Y_i$) | Indpt. Variables | B | B | B | B | B |
| LOGMNB | Intercept | 1.707* | 1.638 | 1.528* | 1.338* | 1.887* |
| | FLR | -0.007* | -0.005 | -0.002 | -0.003 | -0.001 |
| | PURBAN | | -0.001 | 0.002** | -0.002* | -0.002** |
| | BPL | | | 0.001 | 0.003* | 0.002** |
| | GDDP | | | -0.00001066* | | |
| | HINDU | | | | | -0.003 |
| | MUSLIM | | | | 0.009* | 0.005 |
| | CMI | | | | | -0.005* |
| LOGIMR | Intercept | 4.698* | 4.693 | 4.173* | 4.151* | 6.087* |
| | FLR | -0.012* | -0.012 | -0.007 | -0.011** | -0.011** |
| | PURBAN | | -6.27E-05 | 0.005** | 0.001 | 0.002 |
| | BPL | | | 0.01* | 0.011* | 0.008* |
| | GDDP | | | -0.00001398** | | |
| | HINDU | | | | | -0.016* |
| | MUSLIM | | | | 0.002 | -0.018 |
| | CMI | | | | | -0.005 |
| FWPR | Intercept | 71.65* | 50.901 | 50.138* | 56.71* | 67.334* |
| | FLR | -0.586* | -0.122 | -0.16** | -0.189* | -0.24* |

| PURBAN | | -0.289* | -0.321* | -0.235* | -0.244* |
|---|---|---|---|---|---|
| BPL | | | 0.031 | -0.003 | -0.015 |
| GDDP | | | 0 | | |
| HINDU | | | | | -0.119 |
| MUSLIM | | | | -0.313* | -0.461* |
| CMI | | | | | 0.064 |

\* 1 percent LOS, \*\*5 percent LOS

 Source: Derived with the help of SPSS

The Models are same as defined in Table 5.3a

The Linear Regression Equations for the dependent variables are given below Table 5.3b

The Results displayed in the above three tables (Table 5.3a, 5.3b and 5.3c) can be explained as follows:

a) It is observed in the multivariate models (Model 1 to 5, Table 5.3a) that all the independent variables considered have a significant effect in the model. It is found in Model 1 that FLR has a significant effect on the combined multivariate model. From Model 2 to Model 5 I find that PURBAN is having the highest effect in the model (Its Wilks Lambda value is minimum in the model).

b) The Tests of Between-Subjects Effects table of all the Models (Model 1 to 5 of Table 5.3b), gives the significant value of F statistics for all the three Corrected Models of the dependent variables.  By observing the F values and the B estimates in the Table 5.3b and Table 5.3c respectively, I find the factors which have significant effect on the dependent variables. The model wise details are given below.

**Model 1: Design: Intercept + FLR**

From Table 5.3 b and Table 5.3 c I find that the variable FLR has a significant effect on all the three dependent variables (LOGMNB, LOGIMR and FWPR). The adjusted $R^2$ of the three linear regression equations are 0.369, 0.187 and 0.462 respectively.

**Model 2: Design: Intercept + FLR + PURBAN**

I again observe that the variable FLR has a significant effect on all the three dependent variables.  PURBAN is significant only on FWPR. The adjusted $R^2$ of the three linear regression equations are 0.380, 0.162 and 0.896 respectively. The value of adjusted $R^2$ in the model for FWPR has increased by about 90 percent as compared to Model 1 i.e. variable PURBAN introduced in this model has a very strong effect on FWPR.

**Model 3: Design: Intercept + FLR + PURBAN +BPL +GDDP**

From the F values and the B estimates in Model 3 of Table 5.3b and Table 5.3c respectively it is observed that the variable PURBAN has a significant effect on all the three dependent variables. FLR is significant only on FWPR. BPL is significant on LOGIMR and Per Capita GDDP (i.e. GDDP) is having a significant effect on LOGMNB and LOGIMR. The adjusted $R^2$ of the three linear regression equations are 0.660, 0.625

and 0.898 respectively. As BPL and GDDP are introduced in this model, the value of adjusted $R^2$ in the ANOVA model of LOGMNB and LOGIMR has increased by 28 percent and 45 percent respectively as compared to Model2. Thus, I can say that the economic factors like BPL and GDDP have a strong effect on LOGIMR and LOGMNB respectively in the models.

### Model 4: Design: Intercept + FLR + PURBAN +BPL + Muslim

In Model 4 it is found that the variable FLR is having a significant effect on LOGIMR and FWPR. PURBAN and MUSLIM factors are significant on LOGMNB and FWPR. BPL is significant on LOGMNB and LOGIMR. The adjusted $R^2$ of the three linear regression equations are 0.607, 0.564 and 0.921 respectively.

### Model 5: Design: Intercept + FLR + PURBAN + BPL + Hindu+ Muslim + CMI

In the results of Model 5 it is observed that FLR has a significant effect on only LOGIMR and FWPR. PURBAN has a significant effect on LOGMNB and FWPR. BPL has a significant effect on LOGMNB and LOGIMR. HINDU has a significant effect on LOGIMR. MUSLIM has a significant effect on FWPR. CMI has a significant effect on LOGMNB. The adjusted $R^2$ of the three linear regression equations has improved to 0.745, 0.657 and 0.923 respectively.

### 6. Conclusion

All the three dependent variables mean number of births, infant mortality rate and female work participation are significantly correlated with each other in Maharashtra. The findings of the analysis indicate that FLR is significantly and negatively correlated with all the three dependent variable. In Principal Component Analysis, three principal components are extracted and all the factors are found to be important. The variables are divided into following three components namely (Literacy and Other Factors), Religion and Poverty.

It has been observed that the factors which are having a significant effect on the dependent variables are female literacy rate, urbanisation, poverty level, per capita gross district domestic product, religion and medical facilities. Urbanisation is found to have the most significant effect on female work participation. Female literacy is having a strong effect on infant mortality rate and female work participation in most of the models. Female literacy rate is also observed to have a negative effect on female work participation rate.

### 7. Suggestions and Policy Implications

➢ It has been noticed that in Maharashtra female literacy rate has a significant and negative correlated to mean number of births. Hence measures should be taken by the State government to increase the female literacy in order to decrease the fertility.

➢ Female Literacy is also found to be significantly and negatively correlated to infant mortality rate. Thus to improve the figures of infant mortality rate, female literacy should improve in the State.

➢ A significant and negative correlation has been observed between female literacy and female work participation rate. This indicates that higher female literacy is not

leading to higher female work participation. One of the reasons for this scenario might be less job opportunities for literate females. The State government should introspect the reasons behind this problem and find solution for the same.

➢ Muslim Population is also playing a significant and negative role in Maharashtra, Bihar and Kerala, by increasing the fertility and decreasing the female work participation. The Government of these States should assess the reasons behind it and take remedial steps in this direction.

➢ In the results of the study it is observed that the economic and medical indices are negatively correlated with mainly mean number of births and infant mortality rate in the selected State. Hence it is very important to improve these indices as it might help us in reducing these factors.

➢ It is found in the study that urbanisation plays a significant and a strong negative role on female work participation in Maharashtra. Therefore the Maharashtra State Government should initiate steps to investigate the reasons of this negative effect and take proper action to reverse the situation.

➢ Poverty is found to be an important factor affecting the fertility and infant mortality rate in Maharashtra, hence measures should taken by the administration to decrease the percentage of people lying below the poverty line in order to decrease the infant mortality rate.

## 8. Limitations of the Study

❖ The study is based on mainly Secondary data collected primarily from Census 2001.

❖ The study has used fertility and other tables from Census 2001 data, as Census 2011 tables for these variables are not yet available for many variables.

**References:**

• Duggal, Ravi, T. R. Dilip and Prashant Raymus (2005); *Health and Healthcare in Maharashtra, A Status Report,* published by CEHAT, Mumbai.

• Freire, Paulo and Donald Macedo. (1987); *Literacy: Reading the word and the world.*, Bergin and Garvey, South Hadley, A.

• Herz, B. and G.B. Sperling (2004); *What Works in Girls' Education: Evidence from the Developing World*, Council on Foreign Relations Press, New York.

• Hill, M. A., and E. M. King (1993); *Women's Education in Developing Countries: An Overview* in E. M. King and M. A. Hill (eds.) Women's Education in Developing Countries, The Johns Hopkins University Press, Baltimore: pp.1-50.

• Lincove, Jane Arnold (2008); Growth, Girls' Education, and Female Labor: A Longitudinal Analysis, *The Journal of Developing Areas*, 41, 2: pp. 45-68.

- Jejeebhoy, Shireen J. (1996) eds; *Women's Education, Autonomy, and Reproductive Behaviour: Experience from Developing Countries*, Oxford University Press, Oxford.

- Mumbai Human Development Report (2009); *Municipal Cooperation of Greater Mumbai*, Oxford University Press, New Delhi.

- National Human Development Report (2001); planningcommission.nic.in /reports/ genrep/nhdrep/nhdch2.pdf accessed on 5th April 2008.

- Pathak, K.B., Ram, F and Singh, B.S. (1988); A New Method of Estimating Infant and Child Mortality from Data on Children Ever Born and Children Surviving, *Janasamkhya*, 6, 2: pp.159-68.

- Rajan Irudaya S, Nair Mohanachandran P., Sheela K.L., Lalitendu Jagatdeb and Nihar Ranjan Mishra (2008); *Infant and Child Mortality in India District Level Estimates*, Population Foundation of India, New Delhi.

- Registrar General of India, (2001); *Census of India*, Ministry of Home Affairs, New Delhi.Schultz, T. P. (2002); Why Governments Should Invest More to Educate Girls, *World Development*, 30, 2: pp. 207-225.

- Summers, L. (1994); *Investing in All the People: Educating Women in Developing Countries*, EDI Working Papers (No. 45), The World Bank, Washington DC.

- World Bank (2012); *World Development Indicators*, World Bank Washington, DC. http://data.worldbank.org/indicator/SE.SEC.ENRR accessed on 7th September 2012.