Chapter 8

8.0 To identify factors which determine the agricultural land price in a free market sale.

8.1 Literature Review

8.1.1 Extant literature review to identify the variables that were considered for building land valuation models in free market sales.

Appraisal is an ancient art. It has started from the time when men needed to exchange goods. The methodology may be abridged and empiric (Stewart, Land Valuation in Germany, 1937). During early years farm did not have many other alternate uses excepting cultivation. For an agricultural land it is difficult to measure any of its parameters excepting total net returns. Both income and cost figures being empiric, the land valuation has assumed great importance since it has been used for credit and taxation. "Its close connection with the wide range of general economic theory on value and prices has made the discussion on land valuation the fame of something like theory of relativity in the field of agricultural economics" (Brandt, 2014 (Accessed)). David Ricardo brought the concept of economic rent and opportunity cost in the valuation of land. Johann Heinrich von Thunen added the concept of "distance from the market". He posited that closer a piece of land to the urban core the higher would be its market value (economic rent). (Ricardo: Economic Rent and Opportunity Cost, 2017).

In the modern times urban society is extending into a retreating farmland base. New and newer suburbs, shopping centers, utility sites are being developed. This has created new demands. Such demands have outbid farmers for land. The ever expanding urban demand for rural areas affected is the agricultural land in the urban fringe most. The land prices are shooting up. The inflation has been the other major reasons to accelerate land price and its impact is far more than the inflation itself. (Nuckton, 2016). Different plots of agricultural land possess different characteristics. Some of these characteristics such as soil quality and climate cannot be altered. Other features of land such as structural attributes affect the price of land. This relationship is represented by the hedonic price equation P= f (q1, q2, q3…). It is assumed that no individual is able to influence the hedonic price equation. It states that apart from the structural attributes, the population density is also important, although it is not clear whether this is because of the distance to market that has increased the population density or is a result of other factors. (Maddison, A hedonic analysis of agricultural land prices in England and Wales, 2000). C. O' Donoghue et al analyzed the relationship between land use, local markets, environmental and agronomic drivers of land productivity and policy capitalization to understand what drives the farm land market in terms of price making and value of land and to what extent (C.O' Donoghue, 2015). Vantreese et al (1986) has built up agricultural land valuation model which relied on net farm income as the primary determinant of land value. Other variables included voluntary transfers of farmland, government payments, the rate of return on common stock, expected capital gains and the change in average farm size. Model deals with future events, considers specification of expectation functions as a key to the empirical development of the model. Thus, the model requires expectation functions of future rental growth rates, discount rates and inflation rates (Valerie L. Vantreese J. R., 1986).

Comparable sales approach is the most commonly used method for estimating land value when adequate and appropriate data of comparable land sales are available. The land market is not a perfect market. Its value depends on the types of persons and their interest on the particular plot. In reality it varies and varies widely depending on its present and future use. In eminent domain the valuation should be based on highest and most profitable use. This is also true even for valuation in a free sale. In the definition of highest and best use there is a need to check if it is legally permissible, physically possible, and economically viable. There are a few common factors that are considered in computing agricultural land. This includes

- physical characteristics of soil which include attributes like quality of soil, location, fertility and climate;
- social factors like population growth, prestige and education level;

- economic factors like value and income levels, growth and new constructions;
- governmental forces also limit the use of the land through planning and restrictions, taxations etc.

Adjustments are given as % for the variations in the attributes between the reference site and the plots to be assessed. (Gwartney, 1998).

There are other considerations also in valuing agricultural land. Attributes influencing farm land prices can be classified as location, agricultural factor and non-agriculture factors. Results of hedonic regression show that farmland that has higher yield from agricultural production has a higher price, when other things remain constant. When farmland is converted to non-agricultural use there is increase in the farmland price. Prosperous localities attract people in quest for livelihood. This results in in-migration. Population grows. Demand for land increases and hence the price. Land price also increases in localities where farmland is not near any urban center and the land is less likely to be developed immediately for alternative uses. This is because of additional demand arising out of the requirements for non-agricultural use, creates scarcity in the peri-urban area and its induced effect pushes the price up. Land with natural scenic amenities increases land owner's perception of the potential demand for recreational and retirement activities. In such situation price expectations go up. Land located closer to urban areas have better access to urban facilities resulting in higher farmland value (Klaus Drescher, 2001). The agricultural land which is further away will fetch less growth premium. Thus the agricultural land value may be computed as agricultural land rent plus an anticipated growth premium which may be computed as equal to $g/(r^2)$. Thus the value at peri-urban boundary is higher and decays as the distance from the boundary increases. Here g is the future rent and r is common discount rent (Capozza D. R., 1990). And at times the growth premium may easily account for half of the average price of land (Dennis R. Capozza, 1989). Farmland price is generally lower for the larger parcel size (due to demand restrictions), reflecting lower transaction cost for both the buyers and sellers (Chicoine D. L., 1981).

The price of urban land consists of the value of the agricultural land rent and the cost of conversion agricultural land for urban use. The cost should be less than the value of accessibility and an anticipated value an owner can expect to receive as growth premium after development. Capozza also argues “in uncertainty and irreversibility, the rent level that triggers the conversion of land, the reservation rent exceeds the sum of agricultural land rent and the opportunity cost of capital. Analogously, the reservation price of urban land exceeds the capitalized value of agricultural land rents plus the cost of conversion”. Thus it can be concluded that to determine the value of agricultural land, the variable can be classified into two broad categories, namely internal/agricultural variable and the other is external variables.

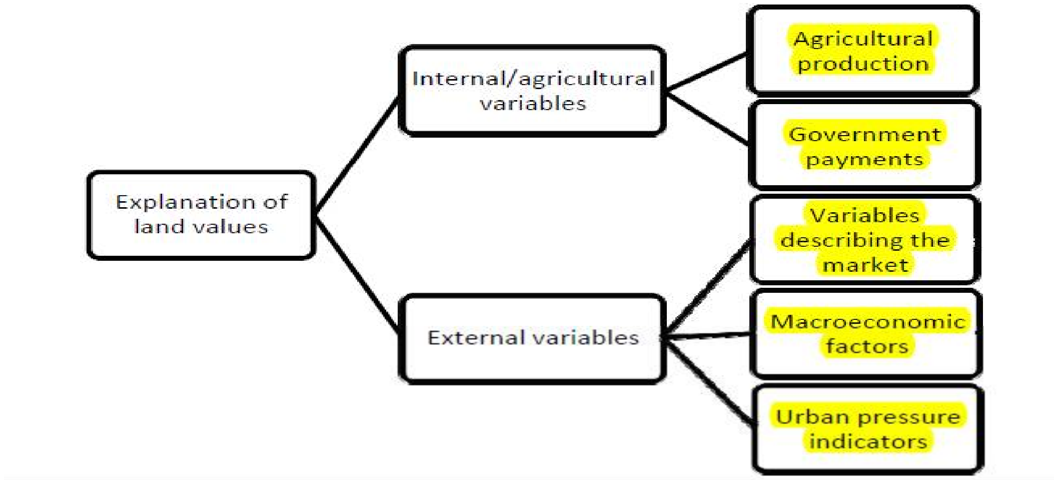
Internal/Agriculture variables are sub-divided into two sub-groups.

a) The first one is related to returns from agriculture production, where soil quality or annual yield can be taken as a measuring parameter.

b) Besides agricultural returns from land, returns from government subsidies are also capitalized into the land price.

Unlike the land rents government subsidies cannot be assumed to be perpetual. It is generally assumed to taper off and can be accounted through a high negative growth rate leading to zero after couple of years. Some of the payments, like agri-environmental payments from government may require additional production costs. In such situations net returns from government payments after setting off the cost may be included. In fact when the government payments are linked with the price of the agriculture produces as percentage, it is difficult to separate rent from agriculture production from government supports and become part of the land value.

Figure 8.1: Variables in empirical analysis



Source: (Paul Feichtinger, 2011)

Feichtinger has identified internal variables as agricultural production and government subsidy. In the subgroup of the external variables he has included variables describing the market, macroeconomic factors and urban pressure indicators. He has further suggested the use of the following measurable parameters which may be used for capitalization.

a) Variables describing the market includes
   a. Farm density
   b. Average farm size
   c. Topography of the land plots
   d. Size of the agriculture land market
b) Macroeconomic factors
   a. Interest rate
   b. Inflation rate
   c. Property Tax rate
   d. Debt to asset ratio
   e. Credit availability etc.

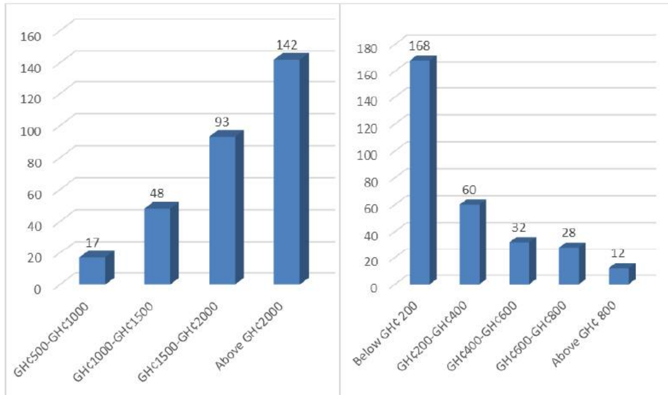
c) Urban pressure indicators

    a. Current population density as measured per square kilometer
    b. Population density as measured in ratio of population to farm acres
    c. Local area population growth
    d. Rurality, measured as fraction of the population's livelihood depends on farms.
    e. Proportion of labor engaged in agriculture etc.

Thus the above discussions can be summarized as follows. To compute the value of agricultural land, the items should be considered for valuation with suitable measuring parameters includes agricultural rent, cost of conversion to non-agricultural use, if any. There is a need to value the accessibility and consequent increase in the future rent, measured as growth premium.

To value the agricultural land and urban land, in External variable the focus is primarily to determine the influence of the distance from the central business district (CBD). But its method of determining the agricultural land prices is ambiguous. Since the agricultural lands are located in rural areas between different CBDs, it is not obvious which CBD should be considered as relevant to determine the price. In case the land is located between multiple urban centers with varying sizes and growth potentials the price estimate needs to be adjusted further. Distance also needs to be adjusted with the accessibility of the urban centers rather than the distance only. In a study conducted in Tamale, Ghana, it was evident from the results that the agricultural land prices had increased significantly when the use change is for residential and commercial purposes. The results are shown in the Figure 8.2 below.

Figure 8.2: Use Change of Agriculture land in Ghana



Land Values Now **Land Values 10 years ago**

Table: Uses of recently acquired land in the study area in Lamale, Ghana

Study has further revealed that it is the type of the area and its current use decides whether and how land values change and how the non-agricultural amenities influence the change as percentage (Bazyli Czyzewski, 2016). It is of interest to note that the lands used for residential purposes had the highest appreciation and that was used in agriculture the lowest. Table below is indicative.

Response	Frequency	Percentage (%)
Agriculture	10	3.3
Commercial	67	22.3
Local Industry	39	13.1
Residential	184	61.3
Total	**300**	**100**

Source: (Francis Zana Naab, 2013)

When land values are calculated using income capitalization approach, it is found that land's income producing value is often unrelated to its market value. Implicit interpretation of the findings can be that price paid for many land purchases cannot be justified economically. There are other forces that are implicit in the land valuation which create divergence. One of the interpretations could be that with land assembly higher productivity can be achieved and the buyer may be willing to pay premium for this. However, this does not explain the increasing interest shown by the non-producers (Helmers, 2004). Land price changes with inflation and study reveals that it increases more rapidly than the inflationary increase in the general price level. Jump in the rate of inflation results an immediate rise in the land price. This may be of interest to note that the rise is often higher. However, after the initial rise, the increase plateaus and stabilizes around the same rate as that of inflation. In his paper Feldstein has considered two portfolio choices to map the impact of inflation. One is for classical financial asset and the other is for land. He has shown that inflation leads to increase in farmland price (Feldstein M. , 1980). In a study conducted in US between 1949 and 1997 on factors affecting land valuation has revealed that "returns from agricultural production are not the key factors in land prices in such areas" (Ian W. Hardie, 2001). Non-farm influences are important (Gardner, 2002) Goodwin et al has shown that land quality also has statistically significant effects along with non-farm influence (Barry K. Goodwin, 2005).

National Council of Real Estate Investment Fiduciaries, USA has considered paired sales approach with the following parameters for adjustments- property rights conveyed, financing terms, condition of sales, date of sale, and also the factors of location, physical characteristics, zoning and any other available characteristics relevant between the two properties (National Council of Real Estate Investment Fiduciaries, 2000). In a study conducted by the Lucas County R&D Section for valuation of residential vacant land using the Spatial Analyst extension of ArcView (Ward et al, 2002) has considered the following variables - "stages" of land development, which is categorized as developing, mature, raw (agricultural) and developing raw. It has also considered land influence which is categorized as ditch, swampy, golf view, wooded etc, and then frontage, lot size, traffic, and followed by location adjustment after the initial model calibration is performed with the earlier factors. In the land valuation the parameters considered includes most profitable use of land and that should be feasible in a shorter horizon of time (Schwenker, 1998). Several physical factors including site specific attributes of the location like number of access points, ease of entry, visibility from roadways, traffic counts, and length of frontage, and location and market factors like supply considerations, future demand factors and competing projects that are planned or under construction. Valuation of land was identified by Pardew as a two-stage procedure where bid (demand) amount which a consumer will be ready to pay for an extra unit for a characteristic and offer (supply) functions which the profit-maximizing firm will accept for a given parcel for individual traits of parcels as one stage. At the second stage the four statistically most significant traits used: parcel size, distance to mountains (scenic object), effective tax rate, and presence of a sewer hookup. Difficulty in estimation arises because of the heterogeneity of the consumers. Different consumers are ready to pay different prices for the same product with the same characteristics (Jolie B. Pardew, 1986). In a study to find factors that influence land prices Winfree has considered current and potential uses of the land, apart from the agricultural revenues, population, and the presence of rivers near the land (Jason A. Winfree, 2002). In a related topic on urban land valuation (Seyfried W. R., 1970) the index of centralization is the location of reference sites. Centrality is typically visualized as a cone of land values, where the cone apex is located in the central business district and the least valued site at the margin of the market. The measure of centrality combines four characteristics: area, value, location, and use.

"Guideline for valuation of immovable properties by the Directorate of Income Tax of India has suggested the use of comparable sales approach. Comparable land should be proximate from time and situation angle. As per the departmental guidelines market value of the comparable lands should be adjusted for i) land size, ii) its shape, iii) total frontage available with the land, iv) location v) nearness to urban amenities and facilities vii) Connectivity and viii) Road width (Directorate of Income Tax Department of India, 2009). In a study to identify factors influencing land valuation in India, the price movement (inflation) in the economy (proxy variable: CPI of Industrial workers of India), land Size, type of the land – agricultural or residential, land within municipal area, geographical factors like distance from rail, road, main district city are found to be most significant (Bhattacharjee et al, 2014). There are non-land factors that cause change in the price expectations. Impact on land price has been mapped using average annual change in population density, "spatial rate of change" in development rents, highway, road traffic density, total farmland acres divided by the county land area, distance from the central business district etc. (Plantinga, 2002) (Shi,et al, 1997).

There is however certain inherent inadequacy in the valuation process. A. Damodaran has commented that there can be no valuation which can be totally objective. The myth and the truth prepared by him is critique's view but these points out some of the probable weaknesses when the assessment is done for valuation (Damodaran, 2006). Damodaran's view is more relevant in the fair market value estimation in eminent domain.

In eminent domain it is an involuntary sale. Study reveals that there is a gap between the giving up price and purchase price. Selling price (often called as Willingness to Accept, WTA) is often much higher than the purchase price, the person is ready to pay (called as Willingness to Pay, WTP). Both are important for fair market value and cannot be ignored (Lewinsohn-Zamir, 2009). Valuing involuntary separation of land (through eminent domain or otherwise) cannot be done through market dynamics of a free sale.

Further, in a thin market where information of comparable sales is less, theoretically income capitalization approach is better suited. But it is difficult to get the income and cost data with acceptable confidence level in agriculture outputs. Prices of the commodity vary depending on when during the year it is sold. This also depends on the location of the agricultural land even for the same crop depending on the plots proximity to road/access to communication. Cost to account for family member's involvement in the field is also difficult. External factors like drought and rain in other region affects demand for agriculture labor and its costs. Income capitalization approach which considers the agricultural yield as the controlling variable thus cannot account for net income accurately. To obviate the weakness comparable sales approach is used to compute net incomes. This however, is an intermittent stage in approximation. In view of this comparable sales approach in valuing land is a preferred option in the Indian context.

8.1.2 The identification of variables which are most suited as attributes to compute fair market value in a thin market to pay compensation.

The above literature review reveals that there are number of factors that affect the agricultural land values; many of them are non-farm factors. There is a need to develop a simple land valuation model which can be logically defended and easily understandable. Only then this can reduce conflict when used to pay compensation for the acquired land.

A closer scrutiny of the above literature review shows that the variables used are mostly case specific. Studies are for active land markets and in a different socio-economic background of the participants in the transactions. The factors need to be reviewed for application in the Indian context. Further, land acquired for development projects through compulsory purchase opens up opportunities for imminent use change to non-agriculture purpose and consequent price rise. Not all projects and in all socio-economic environments will have the same effect on the price. Hence for replacement cost this needs to be factored in. Judicial and legislative guidelines in defining just compensation have specified "highest and best use" of the acquired land. Justice also demands that the economic and social future of the land owners shall be at least as good as it was before acquisition. Not all land owners will be able to migrate to alternate profession after her land is acquired. Hence she should be able to buy a replacement land in the same locality to continue her livelihood. This requires the price computation for compensation to consider the replacement cost rather than the pre-acquisition average sale price. The replacement cost will be equal to the fair market value plus the development induced price changes. To determine the replacement cost of the acquired land it is necessary to identify the variables that are significant in effecting price changes. The variables are noted from the literature review and are listed below. The list is then reviewed and rationalized for use in Indian condition. The variables identified are given below.

Table 8.1: Variables Identified from the Literature Review

Variables Identified from Literature Review			
Sl. No.	**Major Classifications**	**Variables Identified**	**Authors and year**
1	**Agricultural Yield**		
1.1		Soil fertility	Ricardo (1821), Maddison (2000), Gwartney, 1998, Paul Feichtinger, 2011, Barry k. Goodwin, 2005
1.2		Productivity	Ricardo (1821), C. O' Donoghue et al (2015), Yue Jin Shi (1997), Vantreese et al (1986), Klaus Drescher (2001) , Paul Feichtinger, 2011, Ian W. Hardie 2001, Barry k. Goodwin, 2005, Plantinga, 2002, Shi et al, 1997, Jason A. Winfree, 2002

1.3		Climate condition	Maddison (2000), C. O' Donoghue et al (2015), Gwartney, 1998
1.4		Plot size	Chicoine, 1981,Paul Feichtinger, 2011, Helmers, 2004, Directorate of IT Dept. of India, 2009, Bhattacharjee et al, 2014, Ward et al, 2002, Jolie B. Pardew, 1986
1.5		Topography	Paul Feichtinger, 2011, National Council of Real Estate Investment Fiduciaries, 2000, Ward et al, 2002
1.6		Shape	Directorate of IT Dept. of India, 2009
2	**Location of the plot**		
2.1		Distance from the market, Central Business District	Von Thunen (1826,1850), C. O' Donoghue et al (2015), Klaus Drescher (2001), Plantinga, 2002, Ward et al, 2002, Schwenker, 1998
2.2		Distance from urban area	Yue Jin Shi (1997), Gwartney, 1998, Klaus Drescher (2001), Capozza, 1989, Barry k. Goodwin, 2005, Bhattacharjee et al, 2014, Shi et al, 1997, Schwenker, 1998
2.3		Land with natural amenities	Klaus Drescher (2001), Bazyli Czyzewski, 2016, Directorate of IT Dept. of India, 2009, Jolie B. Pardew, 1986, Jason A. Winfree, 2002
2.4		Frontage	Directorate of IT Department of India, 2009, Ward et al, 2002,Schwenker, 1998,
2.5		Connectivity	Directorate of IT Dept. of India, 2009, National Council of Real Estate Investment Fiduciaries, 2000
2.6		Highway/Road width	Directorate of IT Dept. of India, 2009, Plantinga, 2002
2.7		Distance from road/ Distance index	Bhattacharjee et al, 2014, Shi et al, 1997, Jolie B. Pardew, 1986
2.8		Distance from rail station	Bhattacharjee et al, 2014
2.9		Road Traffic Density	Plantinga, 2002, Ward et al, 2002, Jolie B. Pardew, 1986
2.11		Number of access point	Schwenker, 1998
2.12		Ease of entry	Schwenker, 1998

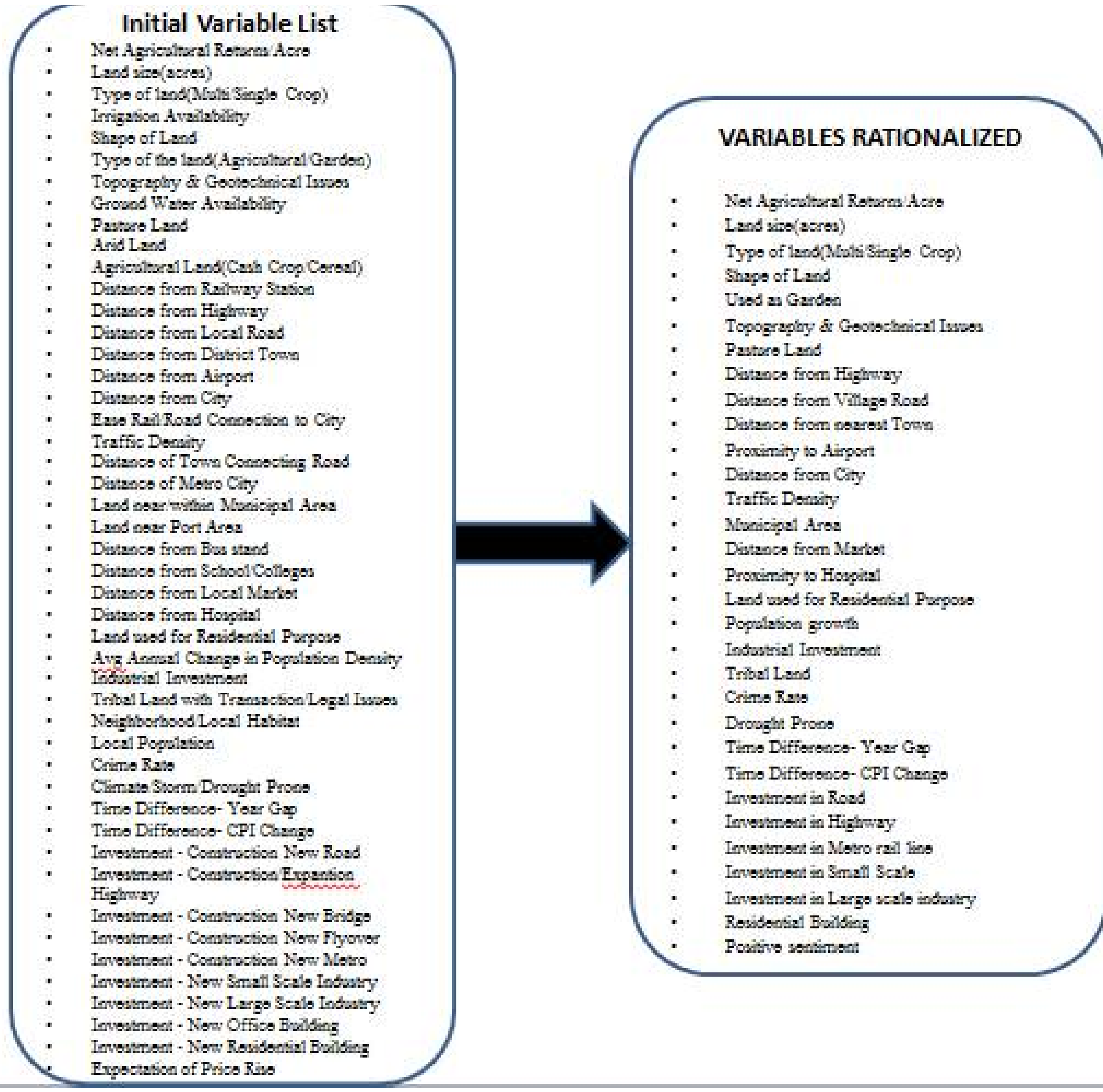
3	**Local Economy and Affluence**		
3.1		Population Density	Maddison (2000), C. O' Donoghue et al (2015), Paul Feichtinger, 2011, Jason A. Winfree, 2002
3.2		Population growth	Yue Jin Shi (1997), Gwartney, 1998, Klaus Drescher (2001), Paul Feichtinger, 2011, Gardner, 2002, Barry k. Goodwin, 2005, Plantinga, 2002, Shi et al, 1997
3.3		Size of the land market	Paul Feichtinger, 2011, Plantinga, 2002
3.4		Debt to Asset Ratio	Paul Feichtinger, 2011
3.5		Proportion of labor engaged in agriculture	Paul Feichtinger, 2011
3.6		Rurality	Paul Feichtinger, 2011, Gardner, 2002
3.7		Non-agricultural amenities	Bazyli Czyzewski, 2016, Directorate of IT Dept. of India, 2009
3.8		Change in development rents	Plantinga, 2002
4	**Time Difference**		
4.1		Inflation	Nuckton (2016), Vantreese et al (1986) , Paul Feichtinger, 2011, Feldstein, 1980, Bhattacharjee et al, 2014
4.2		Date of sale (year gap)	National Council of Real Estate Investment Fiduciaries, 2000
5	**Growth Prospect**		
5.1		Expected capital gain	Vantreese et al (1986), Capozza, 1989, Shi et al, 1997
5.2		Growth and new construction	Gwartney, 1998, Schwenker, 1998
5.3		Prospect of non-agricultural use	Bhattacharjee et al, 2014, Jason A. Winfree, 2002
5.4		Stages of land development	Ward et al, 2002, Jolie B. Pardew, 1986
6	**Government Policy**		
6.1		Policy support	C. O' Donoghue et al (2015)
6.2		Real Interest	Yue Jin Shi (1997), Vantreese et al (1986) , Paul Feichtinger, 2011, Shi et al, 1997

6.3		Government payment	Vantreese et al (1986) , Paul Feichtinger, 2011
6.4		Zoning	Gwartney, 1998, National Council of Real Estate Investment Fiduciaries, 2000
6.5		Taxation	Gwartney, 1998, Paul Feichtinger, 2011, Gardner, 2002, Jolie B. Pardew, 1986
6.6		Credit availability	Paul Feichtinger, 2011
7	**Terms of sales**		
7.1		Property rights conveyed	National Council of Real Estate Investment Fiduciaries, 2000
7.2		Financing (payment) terms	National Council of Real Estate Investment Fiduciaries, 2000
7.3		Condition of sales	National Council of Real Estate Investment Fiduciaries, 2000

Data collected mostly from the western world; the terms and emphasis were more related to the socio-economic environment of the west. These are for free sale in active land market. The list is then taken for discussion with subject matter experts to reframe the variables so that it can be easily understandable by the land owners in general and semi-urban population in particular who are the target respondents in this case. Development induced price changes are rapid in a developing country like India. The variables summarized above were reframed to suit the Indian condition. Thus an initial list of 47 variables was made. This list was further reviewed to remove ambiguities and apparent duplications. Rationalized list of 31 variables was finalized for survey. Variables were ranked in a Likert scale of 1 to 5. The final list of 31 variables along with the initial list of 47 variables is given below. 31 questionnaires in a Likert scale of 1 to 5 were surveyed. The results obtained from the survey were used to identify the principal components, which could account for all the variations that were accounted in the original list of 31variables. The tool used was Factor analysis using SPSS software.

The questionnaire with 31 variables, used in the survey is attached in Annexure-2 (showing a sample electronic response).

Figure 8.3: Variables Rationalized



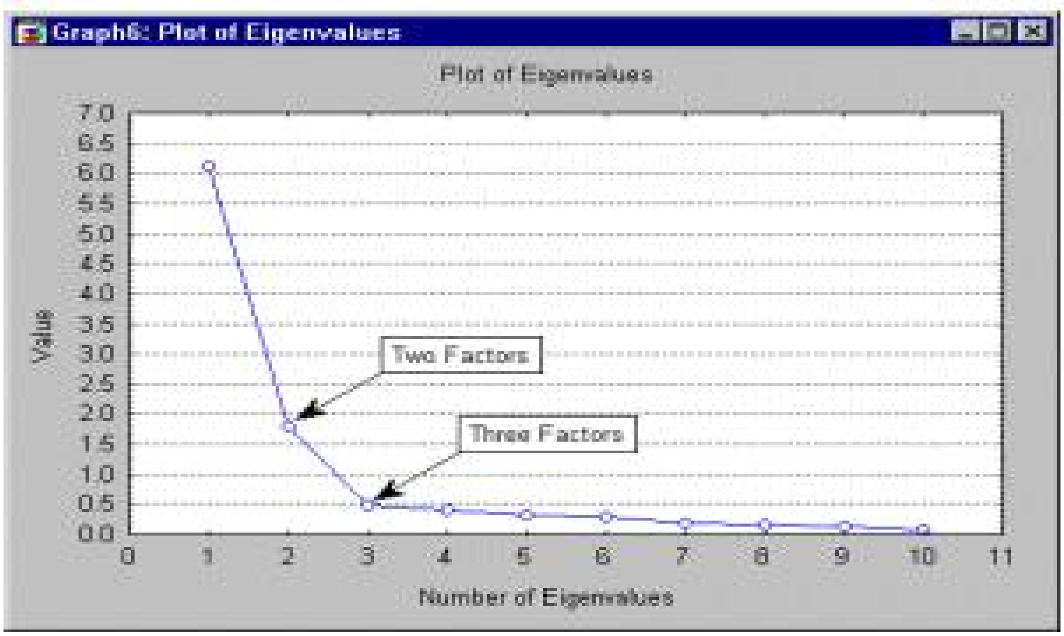
8.1.3: Extant literature review to identify the best suited factor extraction method for principal component analysis

There are number of data reduction and structure detection techniques available with established software supports. Factor analysis may be considered as one of the most popular tools which condenses multiple scaled survey questions down into a fewer number of statistically significant variables. This uses the relationships between variables to classify them into a fewer number of significant variables. This lesser number of variables is then easy to work with and perform other analyses necessary to arrive at decisions with those data. The idea behind factor analysis is that there may be many questions in a survey questionnaire which are interrelated. Factor analysis identifies such questions and the relationships among those interrelated variables through one or more common parameters. This allows condensing of all the related variables down into one new variable or factor. This new variable can then encompass the results of all the original individual variables. Literature review has revealed that the land valuation models are broadly deductive and has considered attributes depending on the location of their study to test the significance. This has made the focus narrow. There is a need for inductive research to identify variables that affect the land price in a country like India with 30 states and almost as many languages and social values and culture. Based on literature review and subject matter expert interviews, 31 variables can be considered as important to effect the land price changes when land is acquired for industry or infrastructure in Indian condition. But it is easier to focus on some key factors rather than having to consider too many variables, all of which may not be very significant. Principal component analysis has thus been used to extract latent variables from the 31 measurable and observable variables to build land valuation model to pay just compensation in eminent domain. Principal Component Analysis is one of the many techniques in data reduction.

Factor Analysis uses Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) techniques depending on the applications. When a set of hypotheses forms the conceptual basis for the factor analysis, a **confirmatory factor analysis** is performed to test the hypothesis. But when there is absence of any guiding hypothesis and the objective is to explore the underlying factors **exploratory factor analysis** is conducted. Exploratory Factor Analysis (EFA) is popularly used statistical tool used in the field of social science. The aim of this is to reveal any latent variables that cause the manifest variable to co-vary (Anna B. Costello, 2005). There are number of factor extraction methods which includes i) Maximum Likelihood method, ii) Principal Axis Factor Method and iii) Principal Component Analysis method. However, Principal Component Analysis itself is a data reduction technique and can be used as a first step and followed by subsequent Factor analysis steps. (An Gie Yong, 2013) . "PCA is a linear dimensionality reduction technique, which identifies orthogonal directions of maximum variances in the original data" (Ranjana Agarwal, 2017). "In PCA, all of the observed variance is analyzed, while in factor analysis it is only the shared variances that are analyzed" (University of Wisconsin, 2010). For data sets where there are many variables, there may be some axes where the variance may be more than the others. In such cases the smaller is generally ignored. This is aimed at reducing the dimensionality of the data set. The objective is to arrive at a few meaningful axes from a larger number of original variables. The name for the approach of rotating data to achieve a smaller number variance is known as Principal Components Analysis, or PCA (Holand, 2008). PCA can be performed either on their variances and co-variances or among the variables and their correlations. In a correlation based PCA the variances of the standardized variables are constructed (W.T.Federer, 1986). This uses Varimax rotation to maximize the variance (variability) of the "new" variable (factor). In spite of this being a data reduction technique, there is no objective guideline to decide how many factors to extract. This rather depends on when there is only very little "random" variability left. Subsequent extraction can only extract less and less variability. However, there are some guidelines available, which are discussed below. This is when used in practice, yields good results.

In the principal component analysis Eigenvectors and Eigenvalues exist in pair. The eigenvalue number is a measure of the spread and the eigenvector with the highest eigenvalue indicates the principal component. Any factor with less than 1 eigenvalue is dropped since this as a factor which has not been able to extract a minimum of one original variable equivalent. Thus the factors with eigenvalues greater than 1 are only retained. This is also known as Kaiser Criteria, and is probably one of the most widely used measures of selecting principal components. There is another popular method where number of factors extracted is decided in a graphical method. This method is known as Scree test. Here the eigenvalues are plotted as shown below and the elbow in a simple line plot is taken as the number of factors.

Figure 8.4: Scree Plot Theory



The decision of when to stop extracting factors remains arbitrary. Using Eigen value sometimes retains too many factors. On the other hand Scree test sometimes retains too few. “Thus it is left to the researcher to choose the one that makes the best sense." (Wherry , 1984).

Ramona Geogescu, et al has studied different data reduction techniques including Partial Least Squares (PLS), Orthogonal Matching Pursuit, etc. apart from PCA. In the most challenging dataset, performing PCA has offered some advantage over other data reduction techniques (Ramona Georgescu, 2017). There are seven techniques of data reduction which was discussed by Rosaria Silipo, et al which includes Missing Values, Low Variance Filter, High Correlation Filter, Random Forests, Backward Feature Elimination, and Forward Feature Construction apart from PCA (Rosaria Silipo, 2014). PCA is the most popular among all of them because of its simplicity in interpretation and understanding. For our use Principal Component Analysis has been used.

8.1.4: To extract latent variables that may be used as factors to build land valuation model in a thin market.

The reliability of factor analysis depends on sample size. For polytomous (with more than two distinct categories) or continuous variables there are two methods to determine sample size. In one the responses are combined and sample size is decided based on their proportions. The other uses the mean. Cochran has developed an equation to decide on the sample size for large population, which is given below (W.G.Cochran, 1963).

$$n_0 = \frac{Z^2 pq}{e^2}$$

where, n_0 is the sample size, Z^2 value is taken from statistical table assuming a normal curve based on the % confidence level (Israel, 1992). Here e is the desired level of precision and p is the estimated proportion of an attribute that is present in the population. q is (1-p).

Determination of sample size- To determine the sample size for the survey, the value for Z is considered for 95% confidence level, i.e. the sampling error of ±5% (precision). This is a large population and it is not known about the variability in proportion. In view of this it is assumed that p=0.5 (maximum variability). With this the resulting sample size will be

$$n_0 = \frac{Z^2 pq}{e^2} = \frac{(1.96)^2\,(.5)(.5)}{(.05)^2} = 385 \text{ sample respondents.}$$

Dr Andy Field has suggested a sample size for large population as 300, when the communalities after extraction comes above 0.5 (Field, 2005). Taro Yamane simplified formula for proportions and has projected the sample size using the following equation (Yamane, 1967)

$$n = \frac{N}{1+N*(e^2)}$$

Here, n is the sample size, N is the population size and e is the acceptable sampling error of say ±5%, the sample size for varying population (abridged version) is given below.

Table 8.2: Sample Size for Precision levels

Sample Size for ±3%, ±5%,±7%, ±10% . Precision level where Confidence level is 95%, and P=0.5				
Size of Population	Sample Size (n) for Precision (e) of			
	± 3%	± 5%	± 7%	± 10%
500	a	222	145	83
1000	a	286	169	91
2000	714	333	185	95
5000	909	370	196	98
10000	1000	385	200	99
50000	1087	397	204	100
100000	1099	398	204	100
>100000	1111	400	204	100

Source- (Israel, 1992).

It may be concluded from the above discussions that a minimum sample size of above 400 should be acceptable for the factor analysis to identify the latent variables that affect the land price in India and should be considered for calculating fair market value of agricultural land in India. 430 samples were collected, which met the sampling size requirements. Respondents were chosen from different states of India keeping some sort of balance among the regions.

Sampling was judgmental to take care of geographical, economic and cultural variations of India. The objective was to identify factors which broadly determined the agricultural land price in India. Data sample and location break-up of the respondents were discussed hereinafter.

8.2. Data Collections

To get a pan India view on land price changes sampling targets included respondents from the hills of Uttarakhand in the north along with Haryana and Uttar Pradesh (UP) to coastal states of Tamil Nadu and Karnataka. Respondents included people from West Bengal, Jharkhand and Bihar in the East to Punjab and Maharashtra in the West. Response of central India at times vary from other regions of India. To make the response give a wider base, respondents from Madhya Pradesh was separately captured as central region.

Care was taken to meet the sampling objective of representative demographic break- up. Here the focus was to get the view of the population who were either directly connected to some sort of economic activities with the agricultural land or might be people who had reasonable knowledge of buying or selling price of land. Land being an attractive investment destination, people working in government offices or in industries were also found to take lot of interest in the price movement of agricultural land. The respondents included this sector of the population also. The demographic break-up of the respondents are given below

Table 8.3: Regional Break-up of the Respondents

Region	States	Respondents	Sub-total
East	West Bengal	150	**180**
	Jharkhand	20	
	Bihar	10	
West	Punjab	30	**40**
	Maharashtra	10	
Central	Madhya Pradesh	50	**50**
North	Uttarakhand	80	**110**
	Haryana	20	
	UP	10	
South	Tamil Nadu	40	**50**
	Karnataka	10	
Total			**430**

Table 8.4: Demographic break-up of the respondents

Location	Occupation	Number	Total
Rural	Farming	60	
	Other Jobs	100	160
Urban/Semi-urban	Government Servants	80	
	Industry employee	60	270
	Real estate allied jobs	40	
	Others	90	
Total			430

8.3 Interpreting Output for Factor Analysis

Result of the factor analysis is interpreted below. The output may be classified into the followings-

Output-1: This shows an abridged version of the R-matrix.

Output-2: This shows the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity.

Output-3: Total Variance Explained the lists of eigenvalues associated with each linear component (factor) before extraction and after extraction.

Output-4: This shows the table of communalities before and after extraction

Output-5: Component Matrix before Rotation

Output-6: Scree Plot, when Kaiser rule may not be accurate

Output-7: Rotated Component Matrix with orthogonal rotation as the factors are independent.

8.3.1 Output-1

31 variables are arranged in R-matrix. Factor analysis strives to reduce this R-matrix down to its underlying dimensions through clustering of variables in a meaningful way. This data reduction is achieved for variables that correlate highly within the group but do not correlate with the variables of the other groups. An R-matrix is a correlation matrix, which is a table of correlation coefficient between variables. As expected the diagonal elements of an R-matrix is all 1.000, since all variable correlate perfectly with itself. The top half of the off-diagonal elements indicates the Pearson correlation coefficient. A closer scrutiny indicates that none of them are greater than 0.90, which indicates that there is no singularity of data, i.e. not measuring the same underlying dimension. Further the determinant is (shown at the bottom of the table) 1.428 E-005. The determinant is greater than 0.00001 and is acceptable. Data is not suffering from multi-collinearity.

crop	Shape	garden	Topography	Pasture	Disthighway	Distvillageroad	Disttown	Distairport	Distcity	Trafficdensity	Municipalarea	Distmarket	Disthospital	Residential	Populationchange	Indinvestment	Triballand	Crimerate	Droughtprone	Yearchange	CPIchange	Invroad	Invhighway	Invmetroline	Inv
47	.196	.129	.072	.137	.409	.043	-.017	.034	.403	.037	.044	.393	.223	.186	.309	.200	.048	.150	.149	.220	.287	.365	.374	.091	
95	.420	.360	.322	.284	.241	.308	.229	.256	.254	.233	.250	.103	.190	.064	.283	.311	.286	.282	.331	.145	.266	.324	.179	.180	
00	.261	.162	.129	.143	.305	.035	.035	.123	.260	.060	.083	.260	.163	.111	.292	.278	.088	.226	.128	.225	.225	.310	.257	.069	
51	1.000	.309	.382	.201	.091	.199	.223	.244	.177	.273	.104	.013	.251	.035	.188	.181	.214	.248	.226	.144	.125	.155	.084	.119	
52	.309	1.000	.430	.163	.020	.133	.304	.233	.057	.145	.202	-.060	.067	-.014	.180	.245	.161	.210	.244	.066	.076	.191	.009	.172	
29	.382	.430	1.000	.346	.003	.312	.468	.464	.145	.323	.204	-.221	.043	-.044	.118	.226	.428	.430	.428	.078	.072	.178	-.056	.132	
43	.201	.163	.346	1.000	.019	.139	.162	.204	.029	.163	.159	-.222	.124	-.094	.077	.016	.211	.226	.315	.026	.034	.035	.016	.029	
05	.091	.020	.003	.019	1.000	.168	.089	.134	.534	.135	.103	.405	.281	.217	.288	.273	.047	.166	.071	.312	.389	.353	.485	.247	
85	.199	.133	.312	.139	.168	1.000	.424	.328	.240	.394	.241	-.005	.040	.103	.105	-.018	.255	.271	.329	-.013	.074	.189	.067	.164	
35	.223	.304	.468	.162	.089	.424	1.000	.380	.240	.354	.346	-.163	.075	.020	.081	.131	.337	.270	.283	.052	.090	.161	.082	.117	
23	.244	.233	.464	.204	.134	.328	.380	1.000	.257	.453	.311	-.130	.253	.145	.204	.227	.402	.477	.268	.081	.103	.256	.118	.289	
50	.177	.057	.145	.029	.534	.240	.240	.257	1.000	.366	.198	.330	.300	.285	.323	.179	.174	.237	.258	.299	.319	.388	.438	.296	
50	.273	.145	.323	.163	.135	.394	.354	.453	.366	1.000	.251	-.050	.274	.152	.252	.097	.431	.314	.319	.152	.230	.168	.190	.298	
33	.104	.202	.204	.159	.103	.241	.346	.311	.198	.251	1.000	.026	.251	.215	.154	.177	.218	.290	.209	.101	.133	.248	.183	.203	
50	.013	-.060	-.221	-.222	.405	-.005	-.163	-.130	.330	-.050	.026	1.000	.236	.334	.152	.148	-.230	-.028	-.133	.257	.322	.218	.325	.136	
53	.251	.067	.043	.124	.281	.040	.075	.253	.300	.274	.251	.236	1.000	.339	.316	.144	.130	.222	.089	.176	.226	.194	.283	.191	
11	.035	-.014	-.044	-.094	.217	.103	.020	.145	.285	.152	.215	.334	.339	1.000	.273	.193	.032	.164	.001	.073	.151	.239	.231	.222	
02	.188	.180	.118	.077	.288	.105	.081	.204	.323	.252	.154	.152	.316	.273	1.000	.318	.265	.262	.238	.232	.279	.319	.287	.165	
78	.181	.245	.226	.016	.273	-.018	.131	.227	.179	.097	.177	.148	.144	.193	.318	1.000	.131	.277	.124	.220	.193	.238	.093	.122	
38	.214	.161	.428	.211	.047	.255	.337	.402	.174	.431	.218	-.230	.130	.032	.265	.131	1.000	.353	.407	.082	.110	.200	.153	.254	
26	.248	.210	.430	.226	.166	.271	.270	.477	.237	.314	.290	-.028	.222	.164	.262	.277	.353	1.000	.393	.234	.250	.294	.110	.208	
28	.226	.244	.428	.315	.071	.329	.283	.268	.258	.319	.209	-.133	.089	.001	.238	.124	.407	.393	1.000	.246	.177	.207	.117	.106	
25	.144	.066	.078	.026	.312	-.013	.052	.081	.299	.152	.101	.257	.176	.073	.232	.220	.082	.234	.246	1.000	.682	.309	.224	.134	
25	.125	.076	.072	.034	.389	.074	.090	.103	.319	.230	.133	.322	.226	.151	.279	.193	.110	.250	.177	.682	1.000	.294	.360	.206	
10	.155	.191	.178	.035	.353	.189	.161	.256	.388	.168	.248	.218	.194	.239	.319	.238	.200	.294	.207	.309	.294	1.000	.512	.309	
57	.084	.009	-.056	.016	.485	.067	.082	.118	.438	.190	.183	.325	.283	.231	.287	.093	.153	.110	.117	.224	.360	.512	1.000	.319	
59	.119	.172	.132	.029	.247	.164	.117	.289	.296	.298	.203	.136	.191	.222	.165	.122	.254	.208	.106	.134	.206	.309	.319	1.000	
21	.387	.235	.460	.221	.078	.338	.390	.406	.274	.396	.241	-.088	.210	.108	.135	.142	.423	.373	.415	.120	.173	.212	.142	.194	
08	.135	.288	.264	.028	.231	.101	.300	.219	.306	.170	.262	.117	.135	.162	.107	.256	.196	.138	.182	.184	.224	.263	.240	.379	
16	.027	.087	.157	-.024	.046	.129	.071	.213	.153	.320	.271	.118	.134	.292	.087	.057	.195	.067	.093	.008	.068	.176	.198	.344	
32	.130	.215	.328	.102	.012	.273	.322	.264	.144	.317	.222	.013	.045	.079	.114	.100	.394	.285	.371	.042	.086	.163	.193	.224	
00	.000	.004	.067	.002	.000	.185	.363	.240	.000	.224	.180	.000	.000	.000	.000	.000	.161	.001	.001	.000	.000	.000	.000	.029	
00	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.016	.000	.094	.000	.000	.000	.000	.000	.001	.000	.000	.000	.000	
	.000	.000	.004	.002	.000	.236	.237	.005	.000	.107	.042	.000	.000	.011	.000	.000	.034	.000	.004	.000	.000	.000	.000	.077	
00		.000	.000	.000	.030	.000	.000	.000	.000	.000	.016	.395	.000	.232	.000	.000	.000	.000	.000	.001	.005	.001	.041	.007	
00	.000		.000	.000	.339	.003	.000	.000	.121	.001	.000	.106	.082	.387	.000	.000	.000	.000	.000	.085	.058	.000	.430	.000	
04	.000	.000		.000	.477	.000	.000	.000	.001	.000	.000	.000	.185	.180	.007	.000	.000	.000	.000	.052	.068	.000	.122	.003	
02	.000	.000	.000		.349	.002	.000	.000	.277	.000	.000	.000	.005	.026	.056	.373	.000	.000	.000	.299	.243	.233	.373	.273	
00	.030	.339	.477	.349		.000	.032	.003	.000	.002	.016	.000	.000	.000	.000	.000	.165	.000	.070	.000	.000	.000	.000	.000	
36	.000	.003	.000	.002	.000		.000	.000	.000	.000	.000	.460	.203	.016	.015	.355	.000	.000	.000	.393	.063	.000	.084	.000	
37	.000	.000	.000	.000	.032	.000		.000	.000	.000	.000	.000	.060	.336	.046	.003	.000	.000	.000	.139	.032	.000	.044	.008	
05	.000	.000	.000	.000	.003	.000	.000		.000	.000	.000	.004	.000	.001	.000	.000	.000	.000	.000	.046	.016	.000	.007	.000	
00	.000	.121	.001	.277	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
07	.000	.001	.000	.000	.002	.000	.000	.000	.000		.000	.150	.000	.001	.000	.022	.000	.000	.000	.001	.000	.000	.000	.000	
42	.016	.000	.000	.000	.016	.000	.000	.000	.000	.000		.299	.000	.000	.001	.000	.000	.000	.000	.018	.003	.000	.000	.000	
00	.395	.106	.000	.000	.000	.460	.000	.004	.000	.150	.299		.000	.000	.001	.001	.000	.278	.003	.000	.000	.000	.000	.002	
00	.000	.082	.185	.005	.000	.203	.060	.000	.000	.000	.000	.000		.000	.000	.001	.003	.000	.032	.000	.000	.000	.000	.000	
11	.232	.387	.180	.026	.000	.016	.336	.001	.000	.001	.000	.000	.000		.000	.000	.257	.000	.488	.065	.001	.000	.000	.000	
00	.000	.000	.007	.056	.000	.015	.046	.000	.000	.000	.001	.001	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000	
00	.000	.000	.000	.373	.000	.355	.003	.000	.000	.022	.000	.001	.001	.000	.000		.003	.000	.005	.000	.000	.000	.027	.006	
34	.000	.000	.000	.000	.165	.000	.000	.000	.000	.000	.000	.000	.003	.257	.000	.003		.000	.000	.046	.011	.000	.001	.000	
00	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.278	.000	.000	.000	.000	.000		.000	.000	.000	.000	.011	.000	
04	.000	.000	.000	.000	.070	.000	.000	.000	.000	.000	.000	.003	.032	.488	.000	.005	.000	.000		.000	.000	.000	.008	.014	
00	.001	.085	.052	.299	.000	.393	.139	.046	.000	.001	.018	.000	.000	.065	.000	.000	.046	.000	.000		.000	.000	.000	.003	
00	.005	.058	.068	.243	.000	.063	.032	.016	.000	.000	.003	.000	.000	.001	.000	.000	.011	.000	.000	.000		.000	.000	.000	
00	.001	.000	.000	.233	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000		.000	.000	
00	.041	.430	.122	.373	.000	.084	.044	.007	.000	.000	.000	.000	.000	.000	.000	.027	.001	.011	.008	.000	.000	.000		.000	
77	.007	.000	.003	.273	.000	.000	.008	.000	.000	.000	.000	.002	.000	.000	.000	.006	.000	.000	.014	.003	.000	.000	.000		
06	.000	.000	.000	.000	.052	.000	.000	.000	.000	.000	.000	.034	.000	.013	.003	.002	.000	.000	.000	.006	.000	.000	.002	.000	
12	.003	.000	.000	.282	.000	.019	.000	.000	.000	.000	.000	.007	.002	.000	.013	.000	.000	.002	.000	.000	.000	.000	.000	.000	
71	.288	.035	.001	.307	.170	.004	.070	.000	.001	.000	.000	.007	.003	.000	.035	.120	.000	.082	.027	.437	.078	.000	.000	.000	
57	.003	.000	.000	.017	.403	.000	.000	.000	.001	.000	.000	.396	.178	.052	.009	.019	.000	.000	.000	.191	.037	.000	.000	.000	

98 E-005

8.3.2 Output 2

Output 2 shows KMO and Bartlett's test results. The KMO statistics close to 1 is an indication of the patterns of correlation being compact and the results would yield distinct and reliable factors. The output in KMO statistics and Bartlett's Sphericity are as follows.

Table 8.6: KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.856
Bartlett's Test of Sphericity	Approx. Chi-Square	4661.479
	df	465
	Sig.	.000

KMO measure of sampling adequacy is 0.856 and is significant. Bartlett's test of Sphericity is 0.000 and is rejected.

8.3.3 Output 3

Table 8.7: Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.190	23.195	23.195	7.190	23.195	23.195	3.415	11.015	11.015
2	3.492	11.264	34.459	3.492	11.264	34.459	3.038	9.800	20.815
3	1.949	6.287	40.745	1.949	6.287	40.745	2.366	7.631	28.446
4	1.361	4.391	45.136	1.361	4.391	45.136	2.249	7.256	35.702
5	1.304	4.205	49.342	1.304	4.205	49.342	2.047	6.603	42.304
6	1.215	3.920	53.262	1.215	3.920	53.262	1.927	6.216	48.520
7	1.072	3.458	56.720	1.072	3.458	56.720	1.862	6.007	54.528
8	1.006	3.246	59.966	1.006	3.246	59.966	1.686	5.438	59.966
9	.957	3.089	63.055						
10	.917	2.958	66.013						
11	.826	2.664	68.678						
12	.802	2.586	71.263						
13	.747	2.411	73.674						
14	.714	2.303	75.977						
15	.680	2.195	78.172						
16	.630	2.034	80.205						
17	.563	1.816	82.021						
18	.558	1.801	83.823						
19	.511	1.648	85.471						
20	.488	1.573	87.044						
21	.471	1.521	88.564						
22	.449	1.447	90.011						
23	.436	1.407	91.418						
24	.405	1.307	92.725						
25	.396	1.278	94.003						
26	.379	1.221	95.224						
27	.347	1.120	96.345						
28	.318	1.024	97.369						
29	.313	1.011	98.380						
30	.281	.905	99.285						
31	.222	.715	100.000						

Extraction Method: Principal Component Analysis

Total variance explained by the first 8 factors with eigenvalues greater than 1, is nearly 60%, whereas subsequent 23 variables explain the balance. After extraction also the percentage of variance for eigenvalues greater than 1 remains same. After rotation the relative importance of the 8 factors has become closer. The first two factors accounted for about 21%. The figures are quite close to other six factors whose eigenvalue is more than 1. This confirms the relative importance of the first few factors equalized. Thus based on eigenvalue greater than 1, there are 8 factors which can explain agricultural land price change.

8.3.4 Output 4

Table 8.8: Communalities

	Initial	Extraction
Agricultural yield/ income	1.000	.710
Land size	1.000	.545
Multi-crop output	1.000	.563
Shape of the land plot	1.000	.659
Agricultural land used for garden	1.000	.586
Topography of land	1.000	.644
Pasture land	1.000	.428
Distance from highway	1.000	.621
Distance from village road	1.000	.672
Distance from nearest town	1.000	.657
Proximity to airport	1.000	.576
Distance from city	1.000	.605
Traffic density on the connecting road	1.000	.608
Land within Municipal area	1.000	.387
Distance from market	1.000	.625
Proximity to hospital	1.000	.656
Residential use of land	1.000	.567
Population growth in local area	1.000	.521
Industrial investment causing land use change	1.000	.642
Tribal land	1.000	.607
Crime rate in the locality	1.000	.564
Drought prone area	1.000	.557
Year change	1.000	.823
CPI change	1.000	.772
Investment in road	1.000	.553
Investment in highway	1.000	.623
Investment in metro rail line	1.000	.455
Investment in small scale industry	1.000	.579
Investment in large scale industry	1.000	.644
Land used for Residential building	1.000	.630
Positive sentiment in land price increase	1.000	.510

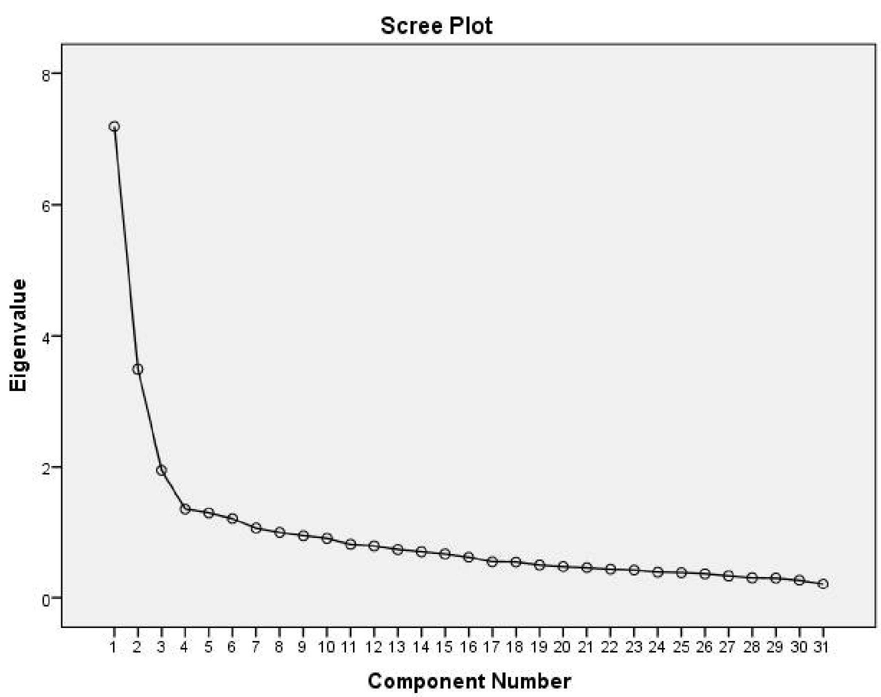
Extraction Method: Principal Component Analysis.

Output 4 in the table of communalities reflect how much variance is truly common. The average communality is 0.5987 and the figures vary between 0.823 (in year change) to 0.387 (in municipal area). Kaiser's criterion suggest that if the average communality is greater than 0.6, then all factors with eigenvalues more than 1 can be retained. However, Kaiser criterion is better suited for lesser number of variables. For the current study the number of variables is 31. There is a need to review the number of factors also in Scree Plot.

8.3.5 Output 5

Factor analysis is an exploratory tool and at this stage the eight factors derived from the SPSS analysis are not taken directly as principal components following the Kaiser criterion. Scree Plots were drawn to decide the number of factors to be extracted and is given below.

Figures 8.5 Scree Plot



The points of inflexion on the curve are shown with arrows. The curve has two tails, once after four factors and then after the seventh. For the purpose of the current study and the average communality being close to 0.60 (Kaiser criterion), the number of factors considered for extraction is taken as seven.

8.3.6 Critical Value

There are varying views of what should be cut off point in suppressing the loadings of each variable onto each factor. Steven's table of critical values against which loadings can be compared, is given below..

Table 8.9: Critical Value

Statistical Significance of Factor Loading	
Sample Size	Critical Values
50	0.722
100	0.512
200	0.364
300	0.298
600	0.21
1000	0.162

The above values are based on an alpha level of 0.01 (two- tailed). In very large samples, smaller loadings can be considered statistically meaningful. Stevens had suggested loadings less than 0.4 as cut off point (Stevens, 2002). However, for the current study, the guideline of the table has been used to choose the cut off value. Loadings less than 0.3 have been suppressed in the output keeping blank spaces in many of the loadings.

8.3.7 Output 6

Based on the above analyses, Component Matrix and the Rotated Component Matrix were run using SPSS for the seven factors to be extracted. Orthogonal rotation was used since the factors are unlikely to be related to each other. In the Rotated Component matrix the first column lists the names of the variables and the second is titled "Component". The sub-columns identify the factors and their loadings. There are some variables which are equally loaded onto more than one factor. The factors where the loading have been the highest were chosen as the factors. As a process loading of one variable onto more than one factor has been avoided. The results of Component matrix and the Rotated Component matrix with 7 factors are given below.

Table 8.10: Component Matrix[a]

	Component						
	1	2	3	4	5	6	7
Agricultural yield/ income	.433	.479	-.340			-.338	
Land size	.589		-.357				
Multi-crop output	.408	.335	-.434				
Shape of the land plot	.466		-.339				
Agricultural land used for garden	.412			.464			
Topography of land	.550	-.525					
Pasture land		-.310	-.361				
Distance from highway	.464	.542					
Distance from village road	.450					-.328	-.428
Distance from nearest town	.494	-.413					-.454
Proximity to airport	.589	-.316					
Distance from city	.601	.344					
Traffic density on the connecting road	.582						
Land within Municipal area	.469						
Distance from market		.710					

Proximity to hospital	.430				.435	.308	
Residential use of land	.322	.340	.381		.320		
Population growth in local area	.497						
Industrial investment causing land use change	.415			.385		.395	
Tribal land	.537	-.383					.323
Crime rate in the locality	.594						
Drought prone area	.542	-.307					
Year change	.399	.361		-.415	-.435	.330	
CPI change	.474	.407		-.392	-.357		
Investment in road	.575						
Investment in highway	.473	.474					
Investment in metro rail line	.465		.415				
Investment in small scale industry	.615	-.339					
Investment in large scale industry	.490				-.477		
Land used for Residential building	.339		.571				.320
Positive sentiment in land price increase	.473	-.301					
Extraction Method: Principal Component Analysis.							
a. 7 components extracted.							

Table 8.11: Rotated Component Matrix[a]

	Component						
	1	2	3	4	5	6	7
Agricultural yield/ income		.815					
Land size	.361	.440			.418		
Multi-crop output		.633					
Shape of the land plot	.369				.388		
Agricultural land used for garden					.702		
Topography of land	.563				.472		
Pasture land	.603						
Distance from highway		.606					.329
Distance from village road				.757			
Distance from nearest town				.696			
Proximity to airport	.391			.364		.427	
Distance from city		.528		.401			
Traffic density on the connecting road	.438			.395		.363	
Land within Municipal area				.328		.347	
Distance from market	- .439	.568					
Proximity to hospital						.686	
Residential use of land						.634	
Population growth in local area		.331				.483	
Industrial investment causing land use change					.638	.306	
Tribal land	.664		.307				
Crime rate in the locality	.463					.367	
Drought prone area	.663						
Year change							.865
CPI change							.793
Investment in road		.491					
Investment in highway		.598	.379				
Investment in metro rail line			.607				
Investment in small scale industry	.527		.319	.308			
Investment in large scale industry			.618		.402		
Land used for Residential building			.731				
Positive sentiment in land price increase	.428		.524				

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 12 iterations.

8.4 Factors Identified

The seven factors have been identified with the following constructs which can well define the 31 variables with measurable parameters. These are given below-

Table 8.12: Factor Analysis Summary

Factor Analysis Summary		
Variables	**Factors affecting agricultural land prices**	
Factor – 1		**Locational Remoteness of the village**
Topography of land		
Pasture land		
Traffic density on the connecting road		
Tribal land		
Crime rate in the locality		
Drought prone area		
Investment in small scale industry		**Measure -** Basics communication and Transport non-availability in the village
Factor 2		**Local Area Affluence**
Agricultural yield/ income		
Land size		
Multi-crop output		
Distance from highway		
Distance from city		
Distance from market		
Investment in road		
Investment in highway		**Measure-** Rural population served by all the amenities in the CD Block
Factor 3		**Investment in Non-agricultural Sector**
Investment in metro rail line		
Investment in large scale industry		
Land used for Residential building		
Positive sentiment in land price increase		**Measure-**Percentage of Population engaged in other jobs in the CD Block
Factor 4		**Plot Location**
Distance from village road		
Distance from nearest town		**Measure-** Distance from Motorable road/ Market/ City
Factor 5		**Non-agricultural use of agricultural land**
Shape of the land plot		
Agricultural land used for garden		
Industrial investment causing land use change		**Measure-** Land used for non-agricultural use to total area of village (x100)
Factor 6		**Population Growth**
Proximity to airport		
Land within Municipal area		
Proximity to hospital		
Residential use of land		
Population growth in local area		**Measure-** Decadal growth in rural population in the CD Block
Factor 7		**Time Difference**
Year change		
CPI change		**Measure-** Year on year on Inflation in %

The factors above have identified the groups or cluster of variables. The factors will be used to construct questionnaires to measure the underlying parameters of the factor. The parameters thus used would define the factors and are to be quantifiable and convenient to use while retaining as much of the original information as possible.

8.5 Factor Interpretations

8.5.1 Factor 1

The variables relates to locational remoteness of the village and its poor economy. Draught prone areas and pasture lands are generally perceived to be less fertile. So is the land with uneven topography. All these make agriculture economy of the village weak. Ease of communication is a basic necessity for alternate economies to grow. Villages with weak communication network cannot bring investments where demand for non-agricultural use of land can grow. This limits the use of land to agriculture. In a weak local economy people migrates out for living. Demand for land does not grow and so also the price. Land acquired for development projects normally leads to new demands to grow but in remotely located villages, this requires long gestation period before outside investors identify opportunities to boost the land market. Outside investors not forthcoming, investments if any are limited to small scale and for the consumption of local people. As a whole the economy remains stagnant. Land acquired for development projects fails to bring prosperity in a short period. This is reflected in the price changes of the agricultural land, which is low. To measure the remoteness of the village "Communication and Transport facilities" available in the village has been considered after conversion of alphabetical rankings of Census 2011 of India to numerical rankings as discussed below-

Census definition of "Yes" is communication being present in location (i.e. within the village), "a" stands for available within 5Kms from the village, "b" for available within 5 to 10Kms from the village and "c" stands for beyond. "Yes", "a" "b" and "c" have been converted to numeric values as 4,3,2, and 1and used to build the proxy index for each village of its remoteness. Since better communication is valued higher, the composite index being higher would mean less remote. Following communication facilities from the Census 2011 was considered for the computation of the shadow index-

Bus service, Railway station, Auto/ Modified Auto, Taxis and Vans, Tractors, Cycle rikshaw, Carts driven by animals, Sea/ River Ferry service, Private Courier facilities, Common Service center- Internet.

8.5.2 Factor 2

The next set of variables that seems to cluster in a meaningful way within the group but do not correlate with the variables outside of that group are the followings- agriculture income, land size, multi-crop output, distance from highway, distance from city, distance from market, investment in road and investment in highway. Presence of these variables in a locality is indicative of affluence. These constitute Factor 2. Farmer will be willing to pay for the present rate of return and the discounted value of the expected future return. (Feichtinger, 2011). Multi-crop production depends on soil quality and irrigation facilities available for the land. These have clear influence on returns from lands. Ricardian model explains the existence of land rents from difference in fertility, or more generally, on land quality. Land of a higher quality generates surpluses over land with lower quality which determines the price differentials (Eric Koomen, 2002). Linkage between farm land values and sector solvency directly impact the economic viability of the farm sector (Gutierezz, 2005). Results of the study conducted by Shi et al, 1997 indicate that "both farm income and urban influences have been important factors affecting the value of farm land". Tracts of rural land that have desirable physical and locational characteristics often sell for much higher prices than neighboring land used for agriculture use. .. Demand for non-agricultural use plays an important role in the variation in rural land prices (J.S. Shonkwiler, 1986). Land located close to the market place and highways increases the prospect of its non-agricultural use. Agriculture lands closer to urban cities may be considered to be in the rural-urban fringe, whose prices depend on the buyer’s assessment of its possible future use and consequential price appreciation. Apart from these as A. Alan Scmidt noted " the greater the percentage change in population growth (of the central city and urbanized area), the greater the percentage appreciation in land values" (Schmid, 1968). Investment in road ways increases connectivity of the locality and helps improving goods and people movements. This aids economic activities and consequent prosperity. Improved affluence in the area causes appreciation of land. Land size has varying impacts on the price of the land depending on the buyer's intended use. "If the size of parcels sold is larger than needed, the additional area adds little or nothing to the utility to the buyer”' (Chicoine D. L., 1981). However, rich peasants in India value larger land plot size higher for better productivity. Here the scale varies from the western world. To account for this the affluence of the CD Block (Community Development Block) was considered while building the proxy indices for the second factor. The second factor is named as Local affluence and to build the proxy indices for the CD Block the percentage of the local population consumes the following amenities (all together) are considered. The list of identified amenities is taken from Census 2011. They are as follows-

Education, Medical, Drinking water, Post office, Telephone, Transport communication, Banks, Agricultural credit societies, and Approach by pucca road, Power supply. In census they are measured in % of population served by each of the amenities. For the proxy index of the CD Block the figure is calculated by multiplying the percentage of population consuming all the amenities.

8.5.3 Factor 3

In an effort to identify parameters that affect the agriculture land price, investments in non-agricultural sector is considered to be a cluster of a group of interrelated variables which explains maximum amount of common variance in measuring the impact of the variables on the agricultural land price. Investment in residential building is an indicator of urbanization and is a strong indicator of use change. When the land available for sale is suited for investment in residential building value appreciation is inevitable. Huge difference in the price of agricultural land takes place (Nguyen Thi Dien, 2011). Investment in large scale industries creates job opportunities to the local people through direct and indirect job creation in industries and also in their ancillaries. The non-agricultural use causes rapid changes in value of land (ibid). Investment in Metro-railway lines reduces communication time from the center of the Metro city (CBD). "Housing and urban growth pressures significantly increase agricultural land values" (Barry Goodwin, 2003). Positive sentiment is created from the prospect of early conversion of agricultural land for non-agricultural use resulting price increase. Proxy index for "Investment in Non-agricultural sector" is calculated based on percentage of population engaged in other than agricultural jobs in the CD Block. The data is sourced from the Census 2011.

8.5.4 Factor 4

Plot location is the fourth factor that can be described in terms of the two variables which have clustered in a group and do not correlate with the variables outside of that group. The variables in this factor relate to physical location of the plot which is viewed from the communication convenience. Connectivity with the Central Business District (CBD) is measured with the distance from the village road and also from the town (here it indicates CBD). Connectivity increases opportunity of trade, and facilitates transportation to the CBD. Nearness to the road increases the price. Land value also changes with the distance from the nearest urban center. Land closer to urban center fetches higher price than the one further away. The structure is typically visualized as a cone of land values (Seyfried W. R., 1970). Proxy index of the plot location is built based on the physical location of the land plot vis-a-vis its proximity to roads or highways with the following parameters –

National/State Highway corner plots and Market adjacent -5, NH, State/ Municipal Road adjacent -4, Subsequent Plots-3, Distant plots -2 , Distant fallow lands-1

These are physical locations of the land and the indices are calculated based on the plot's location from the land survey maps.

8.5.5 Factor 5

Non-agricultural use of agricultural land is the fifth factor that can be identified as an independent cluster of variables where the land is used for primarily non-agricultural use. Land used for garden is a more productive use of the agricultural land in terms of its commercial exploitation. Industrial investment and growth in the neighborhood increases population growth and raises demand for land for non-agricultural use and its price (Richard Ward, 2002). Shape of the land becomes important when there is a prospect of the use of agricultural land for non-agricultural use. "Influences such as corner lots, high traffic volume, unusual shape, unusual topography etc. should be given consideration for possible adjustment" (Deevelopment of Land Value Determinations and Tax Maps, Chapter 4, 2016). Proxy Index for the Non-agricultural use of the agricultural land is taken from the "Percentage of non-agriculture use to the total land in the village". The percentage figure is calculated based on the data available in the Census 2011.

8.5.6 Factor 6

Rural migration to urban growth centers for employment and better livelihood opportunity has been taking place since long. This has reduced pressure on agricultural land in the villages and slowed down the price rise. In some localities the trend is reversed. There is influx of people instead. When rural population grows due to influx from outside demand for both residential and commercial use of land goes up. More demand for residential plots is an indication of reduced migration out of the villages. This increases demand for agricultural land for non-agricultural residential use. "Both office and industrial users pay premiums for lands in neighborhoods with high priced houses". (Peiser, 1987). Proximity to hospital is an indication of nearness to the civic amenities. Nearness of such amenities adds value to the land. Proximity to airports also helps in urbanization and increases the demand. This increases price. Land within municipal area provides improved civic facilities. Urban pressure raises demand for land at the peri-urban boundaries for residential and commercial use and pushes the price up (Capozza D. R., 1990). The urban pressure is measured with "decadal growth rate of rural population"..

8.5.7 Factor 7

Land is a finite asset whose supply cannot be increased with demand. With the growth in population the demand for food and housing is increasing with time. There is more and more demand for other non-agricultural use. Agriculture has to fight with the other demands and the price of agricultural land goes up even in a classical inflation free economy. Data reveals that Global farm land index has risen around 50% to over 1800% between 2002-10 (USDA-Eurostat, 2016). Both the inflation and population growth are linked to relative price of land. (Feldstein M. , 1980). Consumer price index by itself is a good proxy parameter to map the agricultural land price in the emerging economy of India (Subhomoy Bhattacharjee, 2014). 7th Factor which measures the time difference affecting the agricultural land price is measured in terms of CPI (Consumer Price Index) increase between the years. CPI figures are sourced from published CPI indices for India.

8.6 Test of Reliability

Before accepting the results of the Factor analysis and the seven factors there is a need to test the reliability of the scale. Cronbach's alpha test is used to assess the internal consistency of a questionnaire (or survey) that is made up of multiple Likert-type scales and items. The Likert questions in the survey/questionnaire in the current study form a scale of 5. Cronbach's test is to determine if the scale is reliable by establishing whether the items on this questionnaire all reliably measure the same construct. Cronbach's Alpha test was carried out on 430 respondents and the value is based on their scores using SPSS. Results of the tests are given below.-

Scale: ALL VARIABLES

Table 8.13: Case Processing Summary

		N	%
Cases	Valid	430	100.0
	Excluded[a]	0	.0
	Total	430	100.0

a. Listwise deletion based on all variables in the procedure.

Table 8.14: Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.885	.885	31

Cronbach's alpha is **0.885, w**hich indicates a high level of internal consistency for our scale of 5 with this specific sample.

Table 8.15: Item Statistics

	Mean	Std. Deviation	N
Agricultural yield/ income	3.99	.906	430
Land size	3.21	1.047	430
Multi-crop output	3.86	.943	430
Shape of the land plot	3.04	.961	430
Agricultural land used for garden	3.53	.902	430
Topography of land	3.38	.980	430
Pasture land	2.74	.731	430
Distance from highway	3.87	.941	430
Distance from village road	3.30	.903	430
Distance from nearest town	3.69	.797	430
Proximity to airport	2.77	.985	430
Distance from city	3.54	.889	430
Traffic density on the connecting road	2.95	.898	430
Land within Municipal area	3.43	.858	430
Distance from market	3.90	.901	430
Proximity to hospital	3.10	.892	430
Residential use of land	3.39	.919	430
Population growth in local area	3.54	.809	430
Industrial investment causing land use change	3.97	.784	430
Tribal land	3.09	.933	430
Crime rate in the locality	3.24	1.029	430
Drought prone area	3.15	.906	430
Year change	3.69	.815	430
CPI change	3.65	.813	430
Investment in road	3.92	.774	430
Investment in highway	3.90	.857	430
Investment in metro rail line	3.61	.888	430
Investment in small scale industry	3.21	.929	430
Investment in large scale industry	3.92	.779	430
Land used for Residential building	3.57	.830	430
Positive sentiment in land price increase	3.69	.808	430

The **Item-Total Statistics** table presents the "Cronbach's Alpha if Item Deleted" in the final column, as shown below.

Table 8.16: Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Agricultural yield/ income	103.83	159.983	.407	.506	.882
Land size	104.61	155.069	.536	.461	.879
Multi-crop output	103.96	160.294	.375	.382	.883
Shape of the land plot	104.78	159.276	.409	.364	.882
Agricultural land used for garden	104.29	161.241	.353	.338	.883
Topography of land	104.44	157.617	.470	.558	.881
Pasture land	105.08	165.163	.236	.276	.885
Distance from highway	103.95	159.098	.428	.493	.882
Distance from village road	104.53	160.590	.382	.406	.883
Distance from nearest town	104.13	161.082	.416	.444	.882
Proximity to airport	105.05	156.501	.514	.456	.880
Distance from city	104.28	156.963	.556	.513	.879
Traffic density on the connecting road	104.87	157.765	.513	.478	.880
Land within Municipal area	104.39	160.429	.413	.303	.882
Distance from market	103.93	165.314	.173	.483	.887
Proximity to hospital	104.73	160.442	.394	.339	.882
Residential use of land	104.43	162.400	.294	.318	.884
Population growth in local area	104.28	160.219	.452	.340	.881
Industrial investment causing land use change	103.85	162.073	.373	.325	.883
Tribal land	104.73	158.624	.452	.435	.881
Crime rate in the locality	104.58	155.572	.526	.428	.879
Drought prone area	104.67	158.668	.467	.433	.881
Year change	104.13	162.131	.354	.550	.883
CPI change	104.17	160.514	.435	.573	.881
Investment in road	103.90	159.210	.529	.444	.880
Investment in highway	103.92	160.063	.431	.506	.882
Investment in metro rail line	104.21	160.073	.413	.335	.882
Investment in small scale industry	104.61	156.652	.542	.484	.879
Investment in large scale industry	103.90	160.900	.437	.390	.881
Land used for Residential building	104.25	163.199	.295	.348	.884
Positive sentiment in land price increase	104.13	161.146	.406	.380	.882

The results of the Cronbach's Alpha (if Item Deleted) checks the validity of the 31 questionnaire in the survey. Results indicate that any removal will reduce the alpha value.

8.7 Summary of Findings

The above factor analysis discussion is summarized below.

Table 8.17: Summarized Factor Loading

Factor Loading

Variables	Factor Loading	Identified Factors
Tribal land	0.884	Locational Remoteness
Drought prone area	0.863	
Pasture/ Fallow land	0.805	
Topography uneven	0.563	
Investment in small scale only	0.527	
Crime rate high	0.463	
Traffic Density low	0.436	
Agriculture income	0.815	Local Area Affluence
Multi-crop production	0.633	
Distance from Highway	0.606	
Investment in Highway	0.596	
Distance from market	0.566	
Distance from the nearest city	0.528	
Investment in road	0.481	
Plot size	0.44	
Residential Building Projects coming up	0.731	Investment in non-agri-cultural sector
Investment in large scale	0.618	
Investment in Metro-rail line	0.607	
Positive sentiment	0.524	

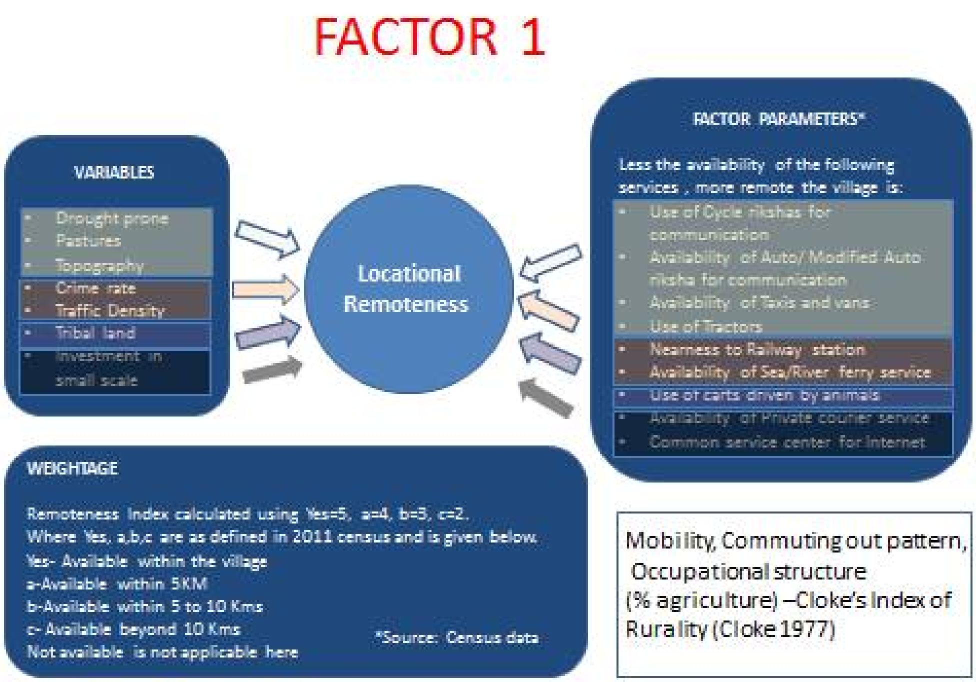
Variables	Factor Loading	Identified Factors
Distance from village road	0.757	Plot Location
Distance from local town	0.696	
Land used for garden	0.702	Non-agricultural use of agriculture land
Industrial investment in the district	0.638	
Shape of the land plot	0.388	
Proximity to hospital	0.686	Population Growth
Close to residential area of the village	0.634	
Population change	0.483	
Proximity to air port	0.427	
Land close to municipal area	0.347	
Year change	0.883	Time change
CPI Change	0.793	

8.8 Identification of observable and measurable parameters which can objectively quantify the identified factors.

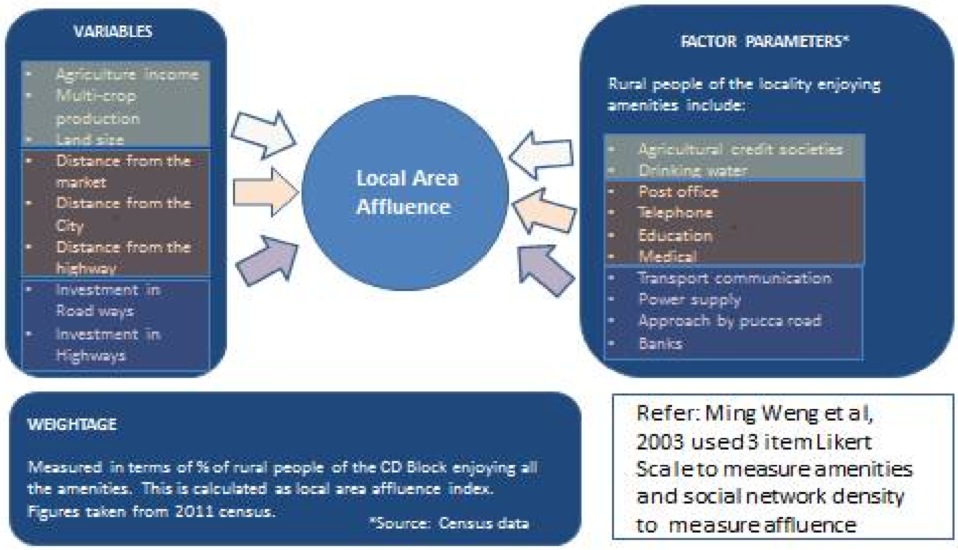
The factors so derived are required to be converted to measurable quantitative parameters so that a rational computation basis can be developed for the valuation model. Census 2011 has identified number of parameters to map the local area development status. These are used as secondary data for the measurement of the 7 identified factors. This requires conversion of the census measures into computable numerical numbers. Communication facilities available, use of electricity, different amenities consumed are mapped in census to measure the growth. Scales used are qualitative and cannot be directly used for building numerical model for land valuation. Prosperity of local area has significant impact on land prices in absolute value and also in their ability to grow. Making best use of the development projects vary depending on the present socio-economic status of the local area. This significantly affects land price increase. The compensation amount by its nature is prospective. In replacement cost the future growth in the land value is necessary to be mapped. All this requires the qualitative measures of census parameters are converted into continuous variable or ordinal numbers. In the summary of factor loading the quantifiable parameters are shown. The numerical conversion in proxy indices is shown in details in the Annexur-4.1 to 4.6..

Summary of Factor Loading and Identified Parameters for Measurements

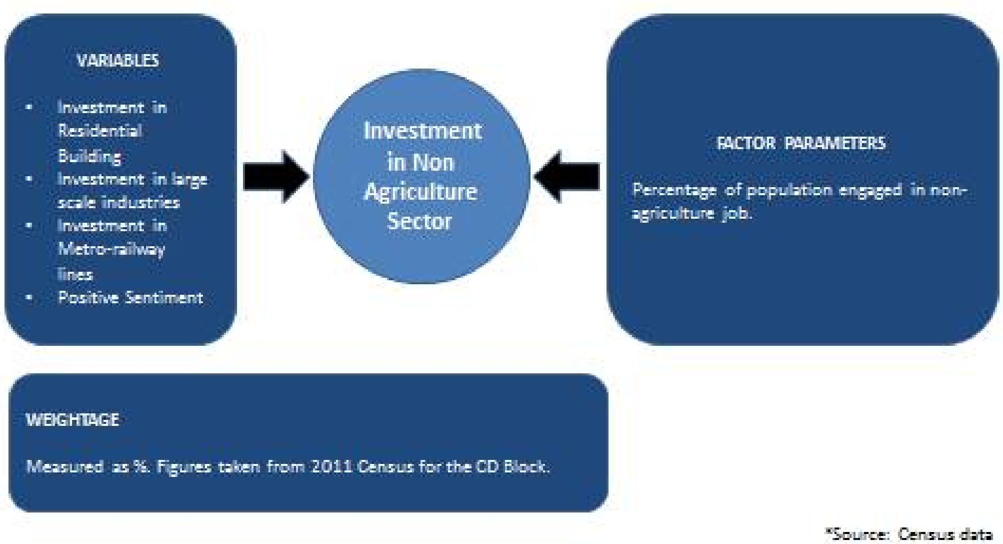
Figures 8.6: Factor 1 to Factor 7



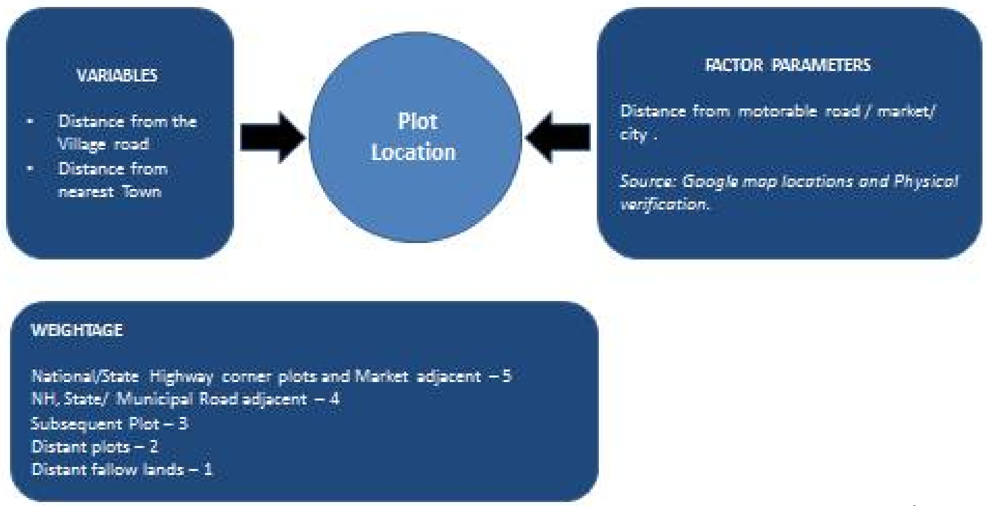
FACTOR 2



FACTOR 3

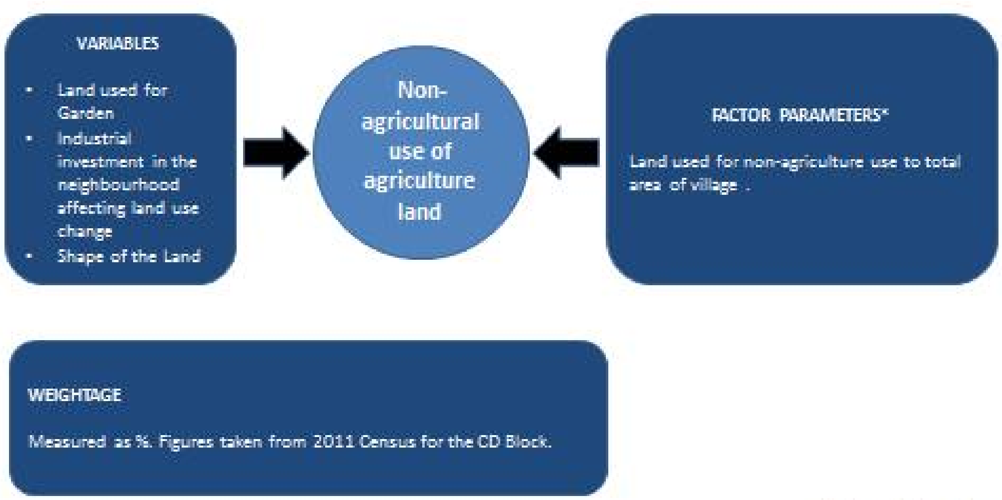


FACTOR 4



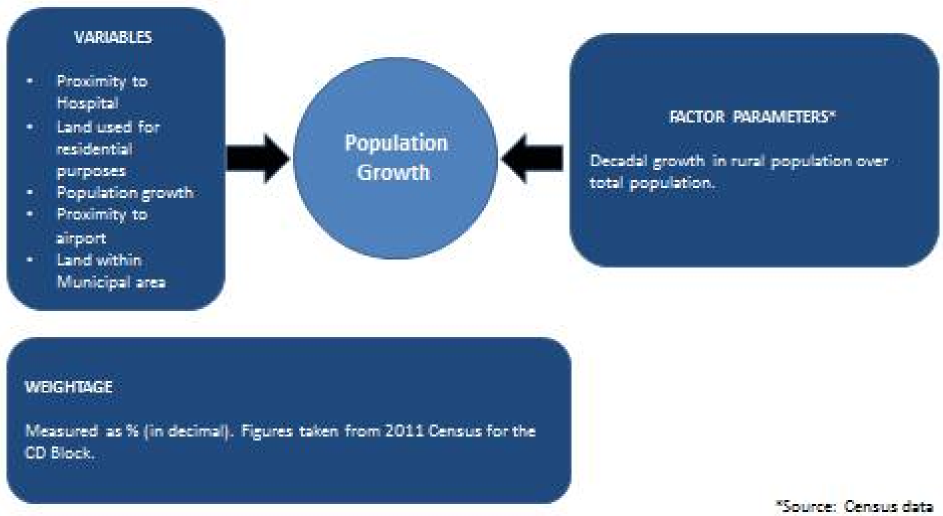
Source: Owner's interview

FACTOR 5

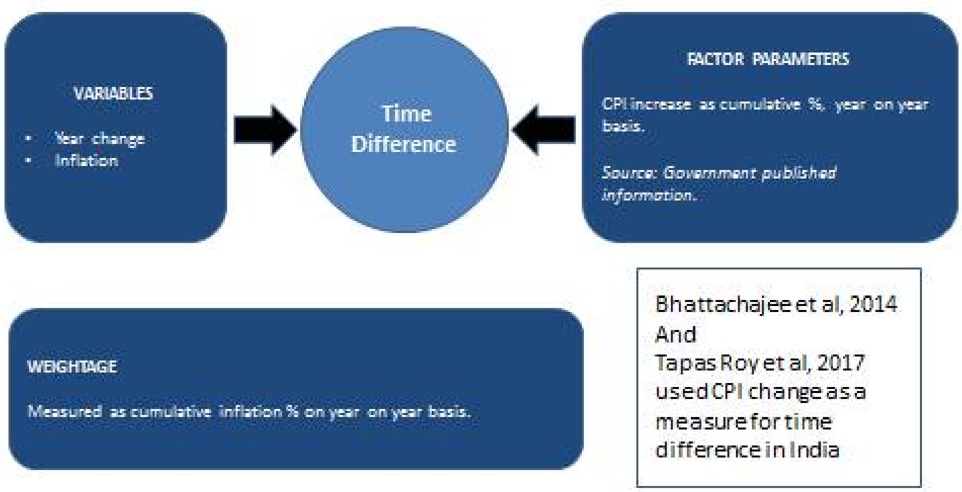


*Source: Census data

FACTOR 6



FACTOR 7



Source: GOI published data