THE DETERMINANTS OF INVESTMENT DISPARITY IN THE ORGANIZED INDUSTRIAL SECTOR OF INDIA

Dissertation submitted to the Jawaharlal Nehru University in partial fulfillment of the requirements for the award of the degree of

MASTER OF PHILOSOPHY

ARPAN KUMAR ROY

Centre for Economic Studies and Planning
School of Social Sciences
Jawaharlal Nehru University
New Delhi 110067
2008

CENTRE FOR ECONOMIC STUDIES & PLANNING
SCHOOL OF SOCIAL SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI- 110 067 (INDIA)

Phone: 91-11-26742575, 26741557, 26742676 Ext. 4421
Direct: 26704421
Cable: JAYENU
Fax: 91-11-26741504, 26741586

CERTIFICATE

Certified that the dissertation entitled **"THE DETERMINANTS OF INVESTMENT DISPARITY IN THE ORGANIZED INDUSTRIAL SECTOR OF INDIA"** submitted by **Arpan Kumar Roy** in partial fulfillment for the award of the degree of **Master of Philosophy (M. Phil.)** of this University, is his original work and may be placed before the examiners for evaluation. This dissertation has not been submitted for the award of any other degree of this University or of any other University.

Prof. Pradipta Chaudhury
(Chairperson)

Centre for Economic Studies
SSS/JNU/New Delhi

Dr. Subrata Guha
(Supervisor)
SUPERVISOR
CESP/SSS/JNU

Acknowledgement

The task of acknowledging the help, moral support, and the immense debt of gratitude I owe to so many people is not an easy task. It gives me hearty pleasure to thank all persons who have helped me in completing this dissertation.

It is not possible to express my gratitude towards my supervisor, Dr. Subrata Guha, for his constant moral support, affectionate encouragement, critical analysis, priceless suggestions at all stages of the dissertation. It would not have been possible for me to complete this exercise without his supervision. I will always consider working under him as a special privilege.

I also express my deep gratitude to all the faculty members, specially Prof. R. P. Sengupta, Prof. C. P. Chandrashekhar and Dr. S. Dasgupta, at the Centre for Economic Studies and Planning for their constant help.

I thank all the staff members of JNU Library for their Constant help. I also owe my gratitude to the staff members of CESP for their kind help.

My seniors and friends have always been extremely cooperative. I gratefully acknowledge the help and encouragement that has been extended towards me by all these special people.

My family has always been supporting in my endeavours. I am indeed grateful to my parents who have always encouraged me in pursuit of higher education. I dedicate the dissertation to them.

Arpan Kumar Roy

THE DETERMINANTS OF INVESTMENT DISPARITY IN THE ORGANIZED INDUSTRIAL SECTOR OF INDIA

Content

Content

Chapter-1

Introduction

1.1 Divergence amongst state in India

One of the main objectives of economic planning in India was to promote balanced regional growth and reduce regional disparity among Indian states by efficient exploitation and effective and balanced utilization of resources. In the 1990s, after four decades of planning experience, economists started reinvestigating how far this objective of economic planning had been achieved. There findings were striking. The results showed that regional disparity had increased among the Indian states despite an increase in the overall growth rate of the Indian economy.

Many of the Indian studies followed the methodology of Barro and Sala-i-Martin (1992), who tested the convergence hypothesis for states in the USA over the period 1840 to 1988. The convergence hypothesis states that because of the diminishing marginal productivity of capital, economies with relatively lower per capita gross domestic product (GDP) should grow at relatively higher rates than their richer counterparts, if the economies, except for their per capita GDP, are similar.Barro and Sala-i-Martin showed that the convergence hypothesis is satisfied for the states of the USA implying that regional disparity among these states decreased.

The trends in regional income disparity can be measured in two ways. Regress growth rate of each economy on the initial per capita GDP. If the estimated coefficient of initial per capita GDP is negative then it can be said that the poorer economies, in terms of initial per capita GDP, are growing at higher rates than the richer ones. This implies that the economies are converging towards their common steady state growth path, reducing inter-economy disparity. This is called β-convergence. The other way is - calculate the coefficient of variation (CV) of per capita GDP for each year under consideration and regress the dispersion on time. If the estimated coefficient of time is negative then it implies that disparity among economies is decreasing over time. This is called σ-convergence.

Ghosh et al (1998) tested the convergence hypothesis for Indian states over the period 1960-61 to 1994-95. They estimated a linear relationship between $(g_i - \bar{g})$ and $(log(x_i) - log(\bar{x}))$ where g_i is the average growth rate of per capita state domestic product(PCSDP) for the ith state, x_i is the initial PCSDP of the ith state and $\bar{g}$ and $\bar{x}$ are the corresponding mean of the two variables respectively. They found a significant positive coefficient of $(\log x_i - \log \bar{x})$ which implies that the richer states are growing at a higher rate. They also calculated the CV of PCSDP for each of the years. This was also found to be increasing, showing increasing disparity over the chosen period.

Rao et al (1999) tested the convergence hypothesis using the same regression model as Barro and Sala-i-Martin (1992) over the period 1965 to 1994. Let us briefly describe the model.

Consider an economy with identical households. Each household has an identical preference pattern. The per capita assets (a) are same for all households and the population grows at an exogenously given constant rate n. The sources of income are wage income by selling labour and income from assets. The wage rate is uniform in the economy. Households can borrow but all borrowing must be repaid with interest. If we assume the number of adults in a household at time 0 is 1, then at time t, the number of adults in a household is given by

$$L(t) = e^{nt} \tag{1}$$

Each household maximizes total utility U which is given by

$$U = \int_0^{\infty} u(c(t))e^{nt}e^{-\rho t}dt \tag{2}$$

Where $c(t)$ is the per capita consumption by an adult at an instant t and let $u(c)$ is given by

$$u(c) = \frac{c^{1-\theta} - 1}{1 - \theta} \tag{3}$$

Now if r be the real rate of interest and if we assume that each individual sells one unit of labour per unit of time, then we can write the change in a households per capita assets per unit of time as

$$\dot{a} = w + ra - c - na \tag{4}$$

From the household utility maximization problem, it can be shown that c would grow according to the following formula

$$\begin{aligned} \frac{\dot{\hat{c}}}{\hat{c}} &= \frac{\dot{c}}{c} - x \\ &= \frac{1}{\theta}[r - \rho] - x \end{aligned} \tag{5}$$

Now let us consider a representative firm in that economy. The production function of the firm is given by

$$Y(t) = F[K(t), \hat{L}(t)] \tag{6}$$

Where $\hat{L}(t)$ is the effective labour, defined as the product of physical labour and the number of efficiency units of labour per physical unit at time t.

$$\hat{L}(t) = L(t)T(t) = L(t)e^{xt} \tag{7}$$

where x is the rate of technological progress in the economy. The marginal products of capital and labour are both positive and diminishing and the production function exhibits constant return to scale in both capital and labour. We can write the production function in the intensive form as

$$\hat{y}(t) = f(\hat{k}(t)) \tag{8}$$

Where $\hat{y} = \frac{Y}{\hat{L}}$ and $\hat{k} = \frac{K}{\hat{L}}$. The marginal products of capital and labour are

$$\frac{\partial Y}{\partial K} = f'(\hat{k}) \tag{9}$$

and

$$\frac{\partial Y}{\partial L} = [f(k) - kf'(\hat{k})]e^{xt} \tag{10}$$

Now a profit maximizing firm would employ factors to its production process to the extent where each factor's marginal product is equal to its price. Therefore,

$$f'(\hat{k}) = r + \delta \tag{11}$$

and

$$[f(k) - kf'(\hat{k})]e^{xt} = w \tag{12}$$

where δ is the rate of depreciation. So substituting the value of w in eq (4), the budget constraint can be written as

$$\dot{\hat{k}} = f(\hat{k}) - \hat{c} - (x + n + \delta)\hat{k} \tag{13}$$

And substituting r in eq (5) the growth of consumption can be written as

$$\frac{\dot{\hat{c}}}{\hat{c}} = \frac{1}{\theta}[f'(\hat{k}) - \delta - \rho - \theta.x] \tag{14}$$

Eq (13) and eq (14) determines the time path of $\hat{k}$ and $\hat{c}$. If we assume that the production function is Cobb-Douglas production function then the log linearization of eq (13) and eq (14) around the steady state gives

$$log(\hat{y}(t)) = e^{\beta.t}log(\hat{y}(0)) + (1 - e^{\beta.t})log(\hat{y^*}) \tag{15}$$

Where $\hat{y}(0)$ and $\hat{y^*}$ are the initial and steady state value of $\hat{y}(t)$ respectively and $\beta(> 0)$ is given by

$$2\beta = [\zeta^2 + 4\frac{(1-\alpha)}{\theta}(\rho + \delta + \theta x)(\frac{(\rho + \delta + \theta x)}{\alpha} - (n + x + \delta))]^{\frac{1}{2}} - \zeta$$

and

$$\zeta = \rho - n - (1 - \theta)x > 0$$

Eq (15) expresses $log(\hat{y}(t))$ as the weighted average of $log(\hat{y}^*)$ and $log(\hat{y}(0))$. Eq (15) also shows that as t increases $\hat{y}(t)$ tends to move toward the steady state value $\hat{y}^*$. β is called the speed of convergence. From eq (15) the average growth rate of $\hat{y}(t)$ over the period 0 and $T \geq 0$ can be written as

$$\frac{1}{T}log(\frac{\hat{y}(T)}{\hat{y}(0)}) = x + \frac{(1 - e^{-\beta T})}{T}log(\frac{\hat{y}^*}{\hat{y}(0)}) \tag{16}$$

Since the economy contains identical households, for the ith economy, over a discrete time period the above equation can be written as

$$log(\frac{y_{it}}{y_{it-1}}) = a_i - (1 - e^{-\beta})[log(y_{it-1}) - x_i(t-1)] \tag{17}$$

where $a_i = x_i + (1 - e^{-\beta})log(\hat{y}_i^*)$ and y_{it} is the per capita output of the economy i at a discrete point of time t. Let us assume that the steady state value $\hat{y_i^*}$ and the rate of technological progress do not vary across economies. So we can say that a_i is constant across economies. Now adding an error term for empirical study, eq (17) can be written, over the discrete period $t - T$ and t, as

$$\frac{1}{T}log(\frac{y_{it}}{y_{i,t-T}}) = a - \frac{(1 - e^{-\beta t})}{T}log(y_{it-T}) + x + u_i \tag{18}$$

Equation (18) expresses growth rate of per capita output of the ith economy, over a discrete time period of length T, as a function of initial per capita output. Since the coefficient of initial per capita output is negative and a and x are constant across economies, if the initial per capita output increases, growth rate would decrease, indicating convergence of per capita output in a group of economies.

Rao et al (1999) found positive and significant coefficient of initial per capita SDP, showing divergence. Rao et al (1999) tested the convergence hypothesis over different sub periods, 1964-94, 1970-94, 1975-94, 1980-94, 1985-94 and 1990-94 and found that disparity has increased over each of the sub periods except 1985-94. For this period the coefficient is positive but not significant. The standard deviation of the per capita SDP was calculated for each of the years and an increasing trend was found over the examined period. To have a deeper insight Rao et al (1999) broke up per capita SDP into three major categories according to sector of origin viz. primary sector, secondary sector and tertiary sector and calculated the standard deviation of per capita SDP separately for all the three sectors. They found that the dispersion in the per capita SDP in the primary sector has been the major source of dispersion of total per capita SDP, followed by the dispersion in the per capita SDP in the secondary sector. But the tertiary sector does not show any consistent trend over the considered period. In order to check the determinants of growth they found that up to 1980 initial per capita SDP had positive and significant impact on economic growth of different states. After 1980, the impact is positive but not significant. The variable "share of loans given by the All India Financial Institutions (AIFI)", which was used as a proxy for private investment, was the most important determinant of growth rate since mid 70s.

Dasgupta et al (2000) obtained the same result as Rao et al (1999) and Ghosh et al (1998). They found that the hypothesis of β-convergence and σ-convergence do not hold for Indian states over the period 1970-71 to 1995-96. They calculated CV of per capita state domestic product across states for each of the years in the concerned period. Then they regressed CV on time and found positive and significant time coefficient, which implied that interstate disparity had increased in the chosen period. To look into the sector wise situation, they broke up per capita SDP into three broad components, viz, SDP in the agricultural sector, in the manufacturing sector and in the tertiary sector and estimated time trend for all the three components as above. The trend was found positive for agricultural and manufacturing sectors but it was negative for the tertiary sector. This result implies that agriculture and manufacturing sectors are responsible for the divergence of per capita SDP. To test the hypothesis of β-convergence, Dasgupta et al (2000) estimated the relationship $ln(Y_{it}) = a + bt$ for each of the state, where Y_{it} is the per capita SDP of the ith state in year t. Collecting the estimated values of b ($\hat{b}$) they regressed $\hat{b}$ on initial per capita SDP (Y_{i0}). Notice that $\hat{b}$ is the growth rate of Y_{it}. They found positive coefficient of Y_{i0}. For rank analysis they ranked 22 major states according to the per capita SDP in descending order for the years 1970-71 to 1995-96 and calculated rank correlation matrix. The significantly high values of correlation coefficients indicated that the rank of different states was stable over the years.

1.2 Objective of the study

Although the central government allots proportionately more fund in favor of the poorer states (poor in terms of per capita net state domestic product)[1]

[1]Ghosh, Marjit and Neogi (1998)

regional disparity has increased over the years. As mentioned above, in 1980s, the disparity in the degree of development of the industrial sector was the second important contributor of the regional disparity in India after the agricultural sector but it was the most important factor in the 1990s. Rao et al's (1999) study indicate that disparity in extent of development of the industrial sector is mainly because of the lack of investment in the poorer states. We have also found increasing investment disparity between states in the organized industrial sector of India. Table-1 showing the result of the regression of CV on a time variable t. The coefficient of the time variable is positive and significant at 5% level of significance. The positive coefficient implies that investment disparity among the states was increasing over the years in the concerned period i.e. some states were getting proportionately more investment than others.

Table 1: Investment Disparity

Independent Variable	Coefficient
t	6.215731* (2.79)
Constant	−12246.34* (−2.77)

Note:dependent variable is CV.

∗ significant at 5% level of significance.

So it is important to identify the factors that influence the regional distribution of investment in India. In this dissertation we have tried to find out the effect of some factors which determine the distribution of investment in the organized industrial sector amongst different states in India over the period 1984 to 2001.

1.3 Structure of the dissertation

The rest of this dissertation is organized as follows. In the first part of chapter-2 we have discussed three investment theories, viz. the neoclassical theory of investment, Kalecki's theory of investment and Tobin's q-theory of investment. In the last part of chapter-2 we have surveyed the existing literature on location choice and distribution of investment. In chapter-3 we have discussed the data and methodology used in this study and in chapter-4 we have presented our results and our conclusions from the study.

Chapter - 2

Factors Affecting Investment and its Distribution - A Survey of the Literature

In this chapter we have discussed three investment theories and surveyed some existing studies on location choice of new investment in order to identify the factors which might determine the distribution of investment across states. In sections 2.1, 2.2 and 2.3 we have outlined the Neoclassical theory of investment, Kalecki's theory of investment and Tobin's q-theory of investment respectively. In section 2.4 we consider the implications of these theories for the *distribution of investment* across regions. In section 2.5, we have discussed the literature on the location choice of new investment and the same has been summarized in section 2.6.

2.1 Neoclassical theory of investment

Let us consider two firms in an economy, one firm say Firm A invests in units of a capital good and rents out units of the capital good and earns rent. The other firm say Firm B rents units of the capital good and uses them in its production process and pays rent (the conclusion would be same if we consider that there is only one firm, involved in both the activities). Now let us consider R to be the market rate of rent at which Firm B rents units of the capital good and P is the price at which it sells its output in the market. So the real cost of renting one unit of the capital goods is R/P. Now the firm's benefit from one extra unit of employment of the capital good is the increase in output - which is called marginal product of capital. Since we assume "diminishing marginal product of capital", as the firm increases its employment of the capital good, the marginal product of capital would decrease. Now to maximize profit, the firm would continue to rent the capital good until the marginal product of capital (mpk) is equal to the rental cost of capital. Therefore

$mpk = R/P$

Now consider Firm A. When the firm invests in a capital good it faces three kinds of cost:

1. If the firm invests taking loan from a financial institution it has to pay interest on the loan. If the firm invests from its savings it loses the same amount that the firm would have earned as interest by depositing in a financial institution. So if P_k is the price of one unit of the capital good and i is the market rate of interest then iP_k is the interest bill that the firm has to pay for investing in one unit of the capital good.

2. When the firm rents out the capital good its value may increase or decrease. If the value of the capital good increases the firm gains and when it decreases the firm loses. So the cost of changes in the value of the capital good is $-\Delta P_k$.

3. Because of the law of nature, capital depreciates over time. If δ be the rate of depreciation,then the cost of depreciation is δP_k.

So the total cost of capital $= (iP_k - \Delta P_k + \delta P_k)$

$$= P_k(i - \tfrac{\Delta P_k}{P_k} + \delta)$$

Now if we assume that the price of the capital good rises with the price of other goods then $\frac{\Delta P_k}{P_k}$ equals the overall rate of inflation. Since interest rate minus inflation rate equals the real interest rate, we can write the cost of capital goods as

$$cost\ of\ capital = P_k(r + \delta) \tag{19}$$

$$\therefore \quad the\ real\ cost\ of\ capital = \frac{P_k}{P}(r + \delta) \tag{20}$$

The profit of Firm A is

$$Profit = Revenue - Cost \tag{21}$$

Now Firm B rents a unit of capital good from Firm A and pays rent which is the revenue of Firm A. So in equilibrium, revenue of Firm A would be equal to the marginal product of capital (in real term) of Firm B.

$$\therefore \quad profit = mpk - \frac{P_k}{P}(r + \delta) \tag{22}$$

So as long as the profit is positive Firm A would continue to invest. So the investment decision function can be written as

$$I = I[mpk - \frac{P_k}{P}(r + \delta)] \tag{23}$$

Eq (23) expresses investment as a function of profit. Since profit is the difference between marginal product of capital goods and cost of capital goods, the factors which affect the *mpk* and cost of capital goods should affect the investment[2].

2.2 Kalecki's theory of investment

According to Kalecki, "investment decisions in a given period, determined by certain factors operating in that period, are followed by actual investment with a time lag. The time lag is largely due to the period of construction, but also reflects such factors as delayed entrepreneurial reactions" [3]. Suppose that, at the beginning of a period, the firms have taken profit maximizing investment decision depending on the constraints and uncertainty that the firm faces in the market where it operates. Now further investment would take place if economic situation changes in the considered period. Kalecki

[2]for neoclassical theory of investment see Macroeconomics (5*th* ed) by Mankiw

[3]Kalecki (1971)

considered three broad category of such changes

(1) changes in gross accumulation of capital by the firms out of current profits, i.e. their current gross savings.

(2) changes in profits.

(3) changes in the stock of fixed capital goods.

The gross savings of the firm is closely related to investment decision. It may be the case that because of the limited capital market, higher amount of investment was not possible at the beginning of the considered period. If the gross savings of the firm increase, it is now possible to undertake higher level of investment and so firm may decide to push its investment level above the previous level. So gross savings is positively related to investment decision.

The other factor that affects investment decision is the increase in profits per unit of time. If profit per unit of time increases, then investment decision may take place in such a project which was thought unprofitable. The existing firms also may decide to invest more to reap more profit in the changed situation. So profit per unit of time is positively related to investment decision.

If the stock of fixed capital goods increases per unit of time and profits remain constant, then the profit rate would fall and the investors would decide to invest less. So the changes in the stock of fixed capital good per unit of time, ceteris paribus, is related negatively to investment decision. So the investment decision D_t taken at time t can be written as

$$D_t = aS_t + b\frac{\Delta P_t}{\Delta t} - c\frac{\Delta K_t}{\delta t} + d \tag{24}$$

Where S_t is the gross savings at time t, $\frac{\Delta P_t}{\Delta t}$ is the increase in profits per unit of time at time t and $\frac{\Delta K_t}{\Delta t}$ is the increment of capital stock per unit of time. a, b, c and d are constant, Following the above discussion, the actual investment can be written as

$$F_{t+\tau} = aS_t + b\frac{\Delta P_t}{\Delta t} - c\frac{\Delta K_t}{\delta t} + d \tag{25}$$

where τ is the length of the time lag. Since change in stock of capital goods is equal to the investment in fixed capital minus the depreciation, we can write eq (25) as

$$\frac{F_{t+\tau} + cF_t}{1+c} = \frac{a}{1+c}S_t + \frac{b}{1+c}\frac{\Delta P_t}{\Delta t} + \frac{c\delta + d}{1+c} \tag{26}$$

The left hand side of eq(26) can be said as the weighted average of the actual investment at time $t + \tau$ and t. Let us assume that the average actual investment is $F_{t+\theta}$, where θ is the length of the time lag , less than τ. Therefore eq (26) can be written as

$$F_{t+\theta} = \frac{a}{1+c}S_t + \frac{b}{1+c}\frac{\Delta P_t}{\Delta t} + \frac{c\delta + d}{1+c} \tag{27}$$

2.3 Tobin's q-theory of investment

Let us consider a firm, which is a representative firm of an industry with N firms. The firm's real profit at time t , neglecting associated purchasing and installation cost of capital goods, is proportional to its stock of capital goods $k(t)$ and a decreasing function of the total stock of capital goods $K(t)$ of the industry. Suppose the functional form of the profit is $\pi(K(t))k(t)$ with $\pi'(.) < 0$. The firm's adjustment cost is a function of the rate of change of firm's stock of capital good $\dot{k}(t)$ which is a convex function with $c(0) = 0$,$c'(0) = 0$ and $c''(.) > 0$,where $c(.)$ is the marginal cost. Let us consider that the per unit purchasing cost of capital good is constant and equal to unity and the depreciation rate is zero and so $\dot{k}(t) = I(t)$ where $I(.)$ is the investment. So the firm's profit at any point of time t can be written as

$$\pi_t = \pi(K(t))k(t) - I(t) - c(I(t)) \tag{28}$$

Now to maximize the present value of its future profits a profit maximizing firm would solve the following maximization problem

$$Max \quad \Pi = \int_{t=0}^{\infty} e^{-rt}[\pi(K(t))k(t) - I(t) - c(I(t))]d(t) \tag{29}$$

subject to $\quad \dot{k}(t) = I(t)$

Therefore the current value Hamiltonian can be written as

$$H(k(t), I(t)) = \pi(K(t))k(t) - I(t) - c(I(t)) + q(t)[I(t) - \dot{k}(t)] \tag{30}$$

The co-state variable $q(t)$ is the shadow price of installed capital good which is the change in profit because of the relaxation of the constraint by one unit. So $q(t)$ is the change in the current value of the future profitability of the firm due to one unit change in the stock of capital goods. So if current value of profits increases by q unit, because of one unit increase in capital goods at time t, then it can be said that the market value of the capital good has increased by q unit. So q is the marginal market value of the firm. The first order condition of the above maximization problem is

$$1 + c'(I(t)) = q(t) \tag{31}$$

The above equation (31) implies that a firm would invest to that point where the cost of one unit more capital good is equal to the marginal value of capital good.

The ratio of the market value of capital goods to its replacement cost is known as *"Tobin's q"*. According to the q-theory of investment the incentives to acquire capital goods depend on the ratio of the market value of capital goods to the replacement cost of capital goods. If the *Tobin's q* is greater than one i.e. if the market value of capital goods exceeds the purchasing and installation cost then the firm would acquire more capital. The more is the value of q the more is the incentive to invest[4].

2.4 Factors that emerge from a survey of investment theories

The three investment theories, discussed above in the beginning of this section, implies various factors that affect distribution of investment. Neoclassical theory of investment indicates the impact of the level of stock of capital goods and rate of interest on the distribution of investment. Because of the diminishing marginal returns of capital, further investment in a region, with higher level of stock of capital goods, may reduce the productivity of capital goods. So the relative size of capital goods has negative influence on the distribution of investment. If output capital ratio assumed same across regions, then theory implies that the regions with higher output in organized industrial sector should have lower investment. As equation (20) implies, the real costs of capital goods increase as the real rate of interest increases. So the real rate of interest has negative influence on the distribution of investment. But if the real rate of interest is same across regions (e.g. across the states in a country like India where interest rate is regulated by the Government of India, and effective at the same rate across all the states of India), then interest rate would have no influence on the regional distribution of investment.

According to Kalecki (1971), investment depends on the level of firm's savings and the change of profit share. Firms often faces the problems of finance because of the limited capital market. So if the level of savings increases, firms

[4]for 'Tobin's q-theory' see Advanced Macroeconomics by David Romer

would get rid of the problems of financing using the savings for investment expenditure. But the level of savings depends on the output level and the the share of profits out of the total output. If the profit share and/or level of output increase, firm's level of profits would increase and so the firm's savings. So the states with relatively greater level of output and/or greater profit margin will have greater investment. Similarly if the rate of change of profit share is greater in a state, then investors would expect greater share of profits in future and so the state should have greater proportion of investment.

Tobin's theory of investment indicates that the investment depends on the market valuation of capital goods. If the replacement costs of capital do not change, then an increase in the market value of capital goods would lead to more investment. So if the market valuation of a capital good in a state is greater than other states with same capital good, then investment would be greater in the the former state. However, it is difficult to obtain data on market valuation of same capital goods located in different states, but it could be assumed to depend on the relative values across state of variables like labour costs, level of infrastructure and per capita income which would determine the prospective returns from capital goods expected by the market.

2.5 Studies of location choice for new investment

The investment theories, discussed above, indicate the variables that determine of the quantum of investment within an economy in both the existing projects and new projects. It does not take into account factors which influence choice of location for new investments within an economy - an issue closely related to the distribution of investment within an economy. To consider this factor we review some studies of location choice in the existing literature. Most of the empirical studies, concerned with the location choice problem, have used conditional logit model as the regression model. Let us briefly outline the model.

Suppose that there are n alternative locations where any firm can invest and the firms are free to choose any location. Let us assume that the profit of a firm depends on the locational characteristics of that location where the firm is situated and the profit function of a representative firm is given by

$$\pi_j = \beta X_j + \varepsilon_j \tag{32}$$

where X_j is the vector of locational characteristics and β is the vector of coefficients and ε is the error term which also captures unobserved characteristics of the location. A profit maximizing firm would choose location j if the profit from jth location is greater than the profit from other alternative locations k i.e. $\pi_j > \pi_k$, for all $k = 1, \cdots, n$, and $j \neq k$. Now if the disturbance term in the profit function is identically and independently distributed and follows the Weibull distribution, then the probability that firm i would choose jth location is given by

$$\begin{aligned} p_{ij} &= prob(\pi_{ij} > \pi_{ik}) \\ &= \frac{e^{\beta X_j}}{\sum_{k=1}^{n} e^{\beta X_k}} \end{aligned} \tag{33}$$

where π_{ij} and π_{ik} are profit of Firm i if it would have been located at location i and j respectively. Now using maximum likelihood estimation procedure we can estimate the above model. Since p_{ij} is the probability of ith firm's investment in the jth location, in this method the amount of investment does not matter. the sign of the estimated coefficient tells us the direction of changes of probability of further investment in the jth location if any of the independent variables in jth location changes, *ceteris paribus*.

Bartik (1985) examined the location choice problem analyzing the location choice by new manufacturing firms in the United States using conditional logit model. He used land area as a proxy for number of potential sites in a state. Since more potential site implies more space for future expansion, the probability of having more investment will be greater than the other state with less potential sites, *ceteris paribus*. So according to Bartik (1985) a state's land area should affect location choice decision positively. He found positive coefficient for the variable. Effects of two labour market variables, percentage of unionized labour and average manufacture wage rate were found significantly negative. The effect of existing manufacturing activity, defined as total manufacturing hours in state per square mile, which also reflects the demand situation and agglomeration economies, was found positive and significant.

The study of Herrin and Pernia (1987) is based on a survey of 100 firms operating in Philippines. These 100 firms, which consist 50 local and 50 foreign firms, have been selected randomly from the top 1000 firms in Philippines in 1983. They interviewed as higher ranking executive officer of the firm as possible (available). They were given 34 different pre selected factors, commonly believed to influence the choice of industrial location, and asked to differentiate them, according to their importance, under four categories viz. decisive, of major importance, of some importance and unimportant. The factors were scored 4 if decisive, 3 if of major importance, 2 if of some importance and 1 if unimportant. Then the factors were ranked in two ways - (a) the mean score and (b) percentage of firms considering the factor as decisive or of major importance. The factors scoring high both by a and b were considered as crucial in the choice of location. 7 out of 34 factors was found to be crucial in the choice of industrial location. These 7 factors were

1. Suitable plot.
2. Reliable electric power.
3. Telephone/telex.
4. Easy road access.
5. Space for expansion.
6. Close to major customers.
7. Suitable building.

There were 7 other factors, which scored high by a but not by b, were considered as of secondary importance. These 7 factors were

1. Public transport
2. Public water supply
3. Property available for lease
4. Low cost of land
5. Industrial infrastructure
6. Labour with required skills
7. Close to major suppliers

Table 2: Rank of scores for crucial and secondary factors

Factors	Rank of score		
	All firms	Local	Foreign
Crucial factors			
Suitable plot	1	1	4
Reliable electric power	2	2	1
Telephone/telex	3	3	2
Easy road access	4	4	5
Space for expansion	5	5	7
Close to major customers	6	6	3
Suitable building	7	7	6
Secondary factors			
Public transport	1	1	2
Public water supply	2	2	1
Property available for lease	3	4	3
Low cost of land	4	3	
Industrial infrastructure	5	5	
labour with required skills	6	6	5
close to major suppliers	7		4

source : Herrin and Pernia(1987)

These 14 factors with their respective rank are shown in Table 2. The important point is that the crucial factors are same for both the foreign and local firms although the ranking among the set is different. Herrin and Pernia (1987) have concluded that "it should be pointed out that the firms being studied here belong to the top 1000 corporations in the country. Thus the key location determinants identified relate to large firms and not necessarily to small and medium sized business concerns, although some of those location factors are probably relevant as well to the latter class of enterprises."

Coughlin et al (1991) studied the determinant of the location choice of foreign direct investment in USA over the period 1981 to 1983 using conditional logit model as described above. They took state land area as a proxy for number of potential sites for investment and found positive and significant effect. They used two variables for local market demand. One is state per capita income and the other is manufacturing density. Manufacturing density may also be a proxy for agglomeration economics. Both the two variables had positive influence on location choice. To capture the labour market, they considered three variables - average state wage rate of manufacturing worker in 1981, percentage of unionized workers in 1980 and state unemployment rate in 1981. The effect of wage rate was negative. The coefficient of the variable unemployment rate, which captures availability of workers, was found positive. But the surprising result of this study was the effect of percentage of unionized workers, which was positive and significant.

Woodward (1992) examined locational characteristics that affect location choice decision analyzing Japanese manufacturing firms in the USA over the period 1980 to 1989. The investigation was done in two steps by using conditional logit model. First he estimated state specific factors' effects and then county specific factors, because " location decisions typically involve at least two stages: first a region or state is selected, then a local site." Let us first indicate the estimated effect of the state variables. Woodward (1992) took gravity adjusted measurement of market using the following formula.

$$MARKET_j = \sum_k \frac{PI_k}{d_{jk}^2}$$

where $MARKET_j$ is the market available for the firm located in jth state, PI_k is the personal income in the kth state and d_{jk}^2 is the distance of kth state from jth state. The distance of the jth state was taken as half of the average radius of the state. Demand situation of a state was captured by this variables and the estimated coefficient was significantly positive. For the labour market variable, the proportion of unionized work force was considered as a proxy and the sign of the coefficient was found negative and significant. In the case of county specific variables, the proxy for agglomeration economics was the number of existing manufacturing plants in a county, which had positive and significant coefficient. Among the labour market variables, the unemployment rate and productivity had positive coefficients. The coefficient of the variable wage rate was negative but not significant. Effect of the other county specific variables educational attainment and land area were significantly positive.

Cheng and Kwan (2000) studied the determinant of location choice of foreign direct investment (FDI) in China over the period 1985 to 1995. The model, they used, focuses on the stock of FDI (because profitability depends on marginal return to capital which is generally assumed to be decreasing). Let us assume that y and y^* are the stock of FDI and desired stock of FDI respectively and they are related by the following formula.

$$\frac{dln(Y_{it})}{dt} = \alpha(ln(Y_{it}^*) - ln(Y_{it}))$$

$$or, \quad \frac{dY_{it}}{dt} = \alpha Y_{it}(ln(Y_{it}^*) - ln(Y_{it})) \qquad (34)$$

where $0 < \alpha < 1$. Equation (34) implies that the percentage change in FDI stock is proportional to the gap between $ln(Y_{it})$ and $ln(Y_{it}^*)$. The equation also indicates a self-reinforcing effect of Y which decreases as the gap between Y_{it} and Y_{it}^* decreases. So the actual accumulation of FDI can be affected by the self-reinforcing effect as well as the factors that changes desired stock of FDI. Now if we consider discrete time intervals then eq (34) can be rewritten as

$$ln(Y_{it}) - ln(Y_{it-1}) = \alpha(ln(Y_{it}^*) - ln(Y_{it-1})) \tag{35}$$

Now collecting terms and denoting the log values by small letters we can write the above equation as

$$y_{it} = (1-\alpha)y_{it-1} + \alpha y_{it}^* \tag{36}$$

Following the existing literature Cheng and Kwan (2000) postulated that the variables infrastructure, labour quality, wage rate, regional income and policies designed to attract FDI affect the desired stock of FDI and formulated the relationship as

$$y_{it}^* = \pi' x_{it} + \lambda_i + \nu_t + \varepsilon_{it} \tag{37}$$

Where x is the vector of variables, π is the vector of coefficients, λ is the unobserved region specific variable, ν is the time specific unobserved variable and ε is the disturbance term. So eq (36) can be written as

$$y_{it} = (1-\alpha)y_{it-1} + \beta' x_{it} + u_{it} \tag{38}$$

where $\beta = \alpha\pi$ and $u_{it} = \alpha\lambda_i + \alpha\nu_t + \alpha\varepsilon_{it}$ They took per capita income and all roads as the proxy variables for regional income, and infrastructure respectively and found significantly positive effect. Two variables, number of special economic zone (SEZ), lagged by year, and ZONE (as the sum of the three variables, number of open coastal cities, economic and technological zones, and open coastal area) were taken as the proxy for policy variables. The effects of these variables were found positive and significant. The effect of the variable wage rate was significantly negative. But the effect of the variable education which was the proxy variable for labour quality, was found insignificant.

Sun, Tong and Yu (2002) examined the determinants of FDI across 30 provinces in China over the period 1987 to 1998. The determinants they considered were - market demand and market size, agglomeration and infrastructure, labour market and political risk. Sun, Tong and Yu (2002) identified more than one proxy variables for those factors and calculated pair wise correlation matrix. They found high correlation coefficient for some variables and so to avoid the problems of multi-collinearity they took seven variables and excluded other. Those seven variables are GDP as a proxy for market demand, length of railway tracks per square km (RAILWAY) as a proxy for infrastructure, domestic investment per worker (PERWI) as a proxy for degree of industrialization, the ratio of cumulative FDI to cumulative domestic investment (CFDI/CINV) as a proxy for level of foreign investment, number of research engineers, scientists and technicians as a percentage of total employees (RSET) as a proxy for labour quality, average wage rate (WAGE) as a proxy for labour cost and risk ranking by political risk services (RISK) as a proxy for political risk. All the variables, except RISK and

CFDI/CINV, were with one period lag value and the model was specified as

$$\begin{aligned} ln(FDI_{it}) &= \alpha_i + \beta_1 ln(GDP_{it-1}) + \beta_2 ln(PERWI_{it-1}) \\ &+ \beta_3 ln(WAGE_{it-1}) + \beta_4 ln(RSET_{it-1}) + \beta_5 ln(RLWAY_{it-1}) \\ &+ \beta_6 ln(RISK_t) + \beta_7 ln(CFDI/CINV_{it}) + \varepsilon_{it} \end{aligned}$$

The model was estimated using both OLS and GLS procedure. Under the OLS estimation only two variables, GDP and RISK, had significant effect at 1 percent level of significance. The effect of GDP, which was a proxy for market demand, was positive and that of RISK, which was a proxy for political risk, was negative. The effects of other variables were insignificant except RSET and Railway showing positive effects at 10 percent level of significance. But the GLS estimation showed highly significant result for all variables except PERWI which was a proxy for degree of industrialization. The effect of GDP, RSET (proxy for labour quality), Railway (proxy for infrastructure) show positive impact on location choice and the variables Wage (proxy for average wage rate) and RISK have negative impact as expected. But the unexpected result is the effect of CFDI/CINV (proxy for level of foreign investment) which is negative. All these results were significant at one percent level of significance.

In order to examine the reasons behind increasing amount of foreign direct investment inflow in the Central and Eastern European countries, Disdier and Mayer (2004) investigated the determinants of French firms' location choice decision from a choice set of 1843 alternative locations distributed over 19 countries in the period 1980 to 1999. The conditional logit model and the nested logit model were used for the estimation. To capture the agglomeration economics, they used number of French firms located in the country, including the concerned firm, as the proxy variable. They found significantly positive coefficient for the variable. The variable GDP and per capita GDP were used as the proxy for demand situation and development level of the host country respectively. Significantly positive coefficient was found for GDP, but that for per capita GDP was negative. The variable average manufacturing wage rate showed significantly negative coefficient.

Chakravorty's (2003) study focuses on the pattern of industrial location choice in the post reform period in India over the period 1992 to 1998. His estimation is based on district level data and all the industries were classified into six categories viz. All Industry, Heavy Industry, Chemicals, Textiles, Agribusiness and Utilities. Two models - Logit model and OLS or Heckman selection model were used to estimate the effects of explanatory variables in all the six categories. The dependent variable in the study was new investment I_{new}. I_{old}, the total existing investment and I_{lag}, the average number of I_{new} in the neighbor districts were two variables for degree of industrialization. Neighbors of a district, suppose district i, was defined as those whose distance from district i (distance between the centroid) is 150 km or less. The effect of I_{old} was positive and significant in all the models but Chemicals and Textiles in the OLS/Heckman models where as I_{lag} had significantly positive effect in all the logit models, but insignificant effects for all industry,Chemicals, Textiles and Utilities in OLS/Heckman model. He used two variables for capital - industrial credit, defined as the per capita lending to local industry by financial institutions and productivity of capital ($PROD-CAPITAL$) measured as the ratio of the value added to fixed capital for existing industry. The effect of these two variables were not significant in both the two types of models. Among the three labour market variables, effect of productivity of labour, measured as the value added per unit of factory labour, had positive effect but was not significant for all the models.

Population of the district were found positive and significant except Textiles in the logit models but was not significant in the OLS/Heckman models. The size of the available labour force had significantly positive impact in some logit models. He used three variables for the indicator of infrastructure (a) percentage of adult literate population, (b) mortality rate at age 5 years per 1000 live births and (c) physical infrastructure which was calculated according to port, airport within 100 km from the district and national highway passing through it, with equal weight for the three. These variables had little overall significance.

Sanyal and Menon (2005) have studied the effect of labour disputes and some other variables that may influence location choice decision in the context of India over the period 1997 to 1999 using conditional logit model. They divided the area of India under 16 locations clubbing the union territories with the nearest states. In this study four labour market variables were considered - indicators of labour conflicts - the number of strikes and lockouts and percentage of unionized workers, state level amendments for industrial disputes, urban workforce participation rate (for size of the labour force, lagged by 1 year) and literacy rate (for efficiency and productivity of labour). A dummy variable was created for the variable 'the number of strikes and lockouts'. The dummy was 1 if the number of strikes and lockouts exceeds the median of the variable, calculated across the states, in the respective year and 0 otherwise. The variable state level amendments for industrial disputes was measured using Besley and Burgess' (2002) classification of all the related amendments, passed in the period 1949 to 1990, into three category, viz. pro-worker, pro-employer and neutral. They took share of pro-worker amendments as a measure of the variable. The other variables of interest were input cost, growth rate of the industrial GSP, the urban Gini coefficient for per capita consumption expenditure (for market accessibility, lagged by 1 year), EXIM bank funding, road length (for infrastructure level, lagged by 1 year). The number of strikes and lockouts dummy and the variable number of unionized worker had significantly negative influence on location choice. They found significantly negative impact of Gini coefficient. The effect of the variable capturing credit availability was positive and significant but the coefficient of road length was significantly negative, which was unexpected. But the authors justified this result as "states with better infrastructure may have higher real estate costs. Such costs may lower location choice probabilities". The effect of growth rate of industrial GSP was found positive and significant. Other variables were not significant.

2.6 Factors which emerge from studies of location choice

The empirical studies concerned to the problem of location choice indicates various variables that affects the choice decision and their effects on location choice decision. The variable per capita income, which is a proxy for the market demand and size of the market, affect location choice decision positively. The labour market variables, wage rate and percentage of unionized workers has negative influence and that of the productivity of labour is positive. The number of potential sites, proxied by the land area has positive effect. Both of the variables infrastructure and existing industrial activity affects the location choice problem positively.

Chapter-3

Data and Methodology

In this chapter, in section 3.1, we have defined the dependent and independent variables, used in our study. The model specification and the estimation procedure has been discussed in section 3.2.

Our study of the determinants of investment in the organized industrial sectors is close to the studies of Chakravorty (2003) and Sanyal and Menon (2005). Both the studies are focused on the distribution of new investment in the post reform period. Chakravorty (2003) studied the district level investment pattern over a period of seven years from 1992 to 1998. Sanyal and Menon's (2005) study mainly focused on the effects of the labour disputes on the distribution of investment. Sanyal and Menon (2005) covered all the states of India including union territories and the states in the North-East region of India and Jammu and Kashmir, which are special category states, in their study over a very short period of three years from 1997 to 1999. Unlike these two studies we have studied a relatively long period from 1984 to 2001. So our chosen period covers the transition period of India from pre reform era to post reform era. We have studied those fifteen major states which are more or less politically stable and were not union territories since 1984.

3.1 Variables used and the data set

In this study we have studied 15 major states of India over the period 1984 to 2001. These fifteen states are Andhra Pradesh, Assam, Bihar, Gujarat, Haryana, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh and West Bengal. Since 90 percent of the total population of India lived in these 15 states (according to Census of India-2001) and they together produce 87.97 percent of industrial net value added of India in the organized industrial sector (according to ASI-2001-02), our assumption is that the result of the study based on these 15 states would be able to describe the all India scenario correctly. So now onward the total value of any variable would mean total of the 15 states. We have assumed that investors choose a location, for investment from the set of available alternative locations, comparing the locational characteristics of all the alternatives, over a period before investment made. We have assumed that period as the previous four years from the year of investment.

The dependent variable of our study is the investment in the organized industrial sector in a state as a proportion of total investment (*inv*) in the organized industrial sector. We have taken investment as proportion of total investment because we are interested in what determines the distribution of investment across states rather than what determines the quantum of investment. The data source is various issues of Annual Survey of Industries (ASI). We have taken net fixed capital formation as investment. Net fixed capital formation (NFCF) is defined as "the excess of net fixed capital at the end of the accounting year over that at the beginning of the year."[5] The consumer price index for agricultural labourers for various states has been used to deflate money values of the net fixed capital formation.

We have chosen six variables that may affect the investment variable, following the literature and availability of data. These four variables are profit share out of per capita net state domestic industrial product (*profit*), pro-

[5]source: ASI.

portion of net state domestic industrial product (*nsdip*), percentage of labour involved in conflicts (*labour*), road length per square km (*road*), per capita net state domestic product *pcnsdp* and per capita electricity consumption (*elec*).

The degree of industrialization or the level of the 'stock of capital goods' is captured by the variable *nsdip*. Since *nsdip* is measured as a proportion of total net domestic industrial product, it captures the effects of the distribution of existing activity on the distribution of investment. A greater level of *nsdip* in any state implies that the industrial sector in that state is large and the industrial infrastructure is relatively better in the state. So because of the better facilities and benefits of external economics of scale, there may be a tendency of concentration of industrial activity in a particular region. On the other hand, greater level of *nsdip* implies high level of stock of capital goods. So, because of the diminishing marginal returns of capital, marginal product of capital goods may be low and the level of existing activity may be negatively related to net investment. ASI does not provide any data under the heading of net state domestic industrial product but net value added in the organized industrial sectors. We refer to that variable as net state domestic industrial product.

While *nsdip* looks at the effects of distribution in scale of production on distribution of investment, the variable $profit$ capture the effect of differences in profitability of existing activity on the distribution of investment. $profit$ not only an indicator of the expected future profitability of the firm but an indicator of availability of internal funds for investment [Kalecki(1971)]. So the region with relatively greater profit share should have greater share of investment. The expected sign of the coefficient of $profit$ is positive. We have collected the data on profit from ASI.

In setting up new industries, profitability of existing activity may not always be an accurate guide and factors like market size, state of labour market, level of infrastructure may need to be considered in addition.

We have taken only one labour market variable - percentage of workers involved in labour conflicts (*labour*). We have collected data on number of total workers involved in labour conflicts from various issues of Indian Labour Statistics, Indian Labour Year Book and Industrial Disputes in India published by the Ministry of Labour, Government of India and total workers from ASI, and calculated percentage of labour involved in conflicts. Since conflicts are undesirable from the managerial point of view, *labour* should have negative coefficient.

We have taken two infrastructure variables, road length per square kilometer (*road*) and per capita consumption of electricity (*elec*), as the indicators of the level of infrastructure. Data for total road length has been collected from various issues of Basic Road Statistics of India, published by the Ministry of Surface Transport, Government of India. Data for electricity consumption have been collected from 'All India Electricity Statistics' published by the Government of India. These variables should have a positive coefficient, but as the literature indicates, may be negative also, because a region with well developed infrastructure may have higher real estate cost than other region.

Market demand situation is captured by *pcnsdp*. Since in a developing country like India, elasticity of demand for industrial product is greater than unity, if per capita income increases, demand for industrial product would increase. So *pcnsdp* should have positive sign. The data for *pcnsdp* has been collected from Central Statistical Organization (CSO), Government of India. The consumer price index for agricultural labourers for various states has been used to deflate money values of the per capita net state domestic products.

The influence of taxation related to industry and government policy to attract industry, have been discussed in some of the studies [6]. In India, apart from the central government, state governments have periodically announced different types of concession to the industries. Before the initiation of Value Added Tax (VAT) state governments in India engaged in a unholy competition in terms of tax concession. Since different state governments were giving different types of concession to the industry, it is difficult in the Indian context to model the effect on the distribution of investment of taxation and government policy to attract industry. So the study of these factors is beyond the scope of the present study.

3.2 Model specification.

Now let us specify the relationship between dependent and independent variables as

Model-1

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2(x_{2it} - \bar{x}_{2t}) + \beta_3(ln(x_{3it}) - ln(\bar{x}_{3t})) + \beta_4(x_{4it} - \bar{x}_{4t})$$
$$+\beta_5(ln(x_{5it}) - ln(\bar{x}_{5t})) + \beta_6(ln(x_{6it}) - ln(\bar{x}_{6t})) + c_i + T_t + u_{it}$$

where, $y : inv$

$x_1 : nsdip$

$x_2 : profit$

[6] Bartik (1985),Coughlin et al (1991), Woodward (1992).

x_3 : *road*

x_4 : *labour*

x_5 : *pcnsdp*

x_6 : *elec*

$x_{jt} = \frac{1}{15}\sum_{i} x_{jit}, \qquad j = 2, 3, \cdots\cdots, 6$

u : *idiosyncratic errors*

and T_t is the year specific effects which is controlled by year dummy. i represents state and t represents year.

In Model-1, we have taken the deviation of each of the variables, representing locational characteristics, from their respective mean, except *nsdip*. Since our dependent variable is the proportion of total investment in each state, so the deviation from mean i.e. the distance from the average rather than the level would correctly explain the variation of the investment.

Apart from the above mentioned explanatory variables, there are unobservable variables, indicated by c_i in Model-1, that may play an important róle in the distribution of investment across states. As specified c_i are the state specific effects that do not vary with time but may vary with states. The main problem of estimation of panel data models is the presence of unobserved effects c_i which may be a random effects like a random variable or fixed effect like a parameter. If we consider the unobserved effects as a random variable then, putting the c_i in the error term, we can estimate the model consistently and efficiently under the following four assumption.

1: $E(u_{it}|x_{i1}, x_{i2}, \cdots, x_{iT}, c_i) = 0, \qquad t = 1\cdots\cdots, T$ and $x_{it} = (x_{1it}, x_{2it}, \cdots, x_{kit})$

2: $E(c_i|x_{i1}, x_{i2}, \cdots, x_{iT}) = E(c_i) = 0$

3: $rank\ E(X_i'\Omega^{-1}X_i) = K, \qquad \Omega \equiv E(v_i v_i')$ and $v_{it} = c_i + u_{it}$

4: $E(u_i u_i'|x_i, c_i) = \sigma_u^2 I_T$

5: $E(c_i c_i' | x_i) = \sigma_c^2$

Assumption 1 and 2 implies the zero correlation between the regressors and the idiosyncratic error term and between regressors and c_i. Assumption 3 implies that the regressors should not be perfectly collinear. Assumptions 4 and 5 are the homoscedasticity assumptions of the combined error term $c_i + u_{it}$. But if we consider the unobserved effects as fixed effects then we can estimate the model even dropping assumption 2 i.e. we can estimate the model under assumption 1, 3 and 4[7] even when the regressors are correlated with the unobserved effects c_i. So fixed effect estimators are efficient and consistent estimators under a weaker set of assumptions than random effect models.

But it is often seen in panel data that the error terms are correlated with the regressors of the other observations within the same unit variables i.e. v_{it} $(= c_i + u_{it})$ is correlated with x_{is} where $t \neq s$. Since assumptions 1 and 2 in the random effect and fixed effect models do not allow such correlation, we can not estimate such a model with both the random effect and fixed effect model procedure. In such a case, we can estimate the model by running an OLS regression with 'robust standard error' after controlling the unobserved effects by dummy variables and clustering the sample with each unit variable as a cluster. So cluster sample OLS regression with robust standard error would produce efficient and consistent estimators under the following two assumptions:

6: $E(v_{it} | x_{it}) = 0$

7: $rank\ E(X_i' X_i) = K$

[7]No need of assumption 5, because in the fixed effect model we consider c_i as parameters.

Before estimation we have calculated pairwise correlation coefficient for all the regressors. The result is shown in Table-3 . Since all the coefficients are less than 0.8, we should not face any problem of multi-collinearity if we use all the regressors in the same model.

Table 3: pair wise correlation coefficient

	nsdip	profit	road	labour	pcnsdp	elec
nsdip	1.0000					
profit	0.0618	1.0000				
road	-0.1050	0.0967	1.0000			
labour	-0.1677	-0.2706	0.0658	1.0000		
pcnsdp	0.3641	0.0633	0.0232	-0.2514	1.0000	
elec	0.4114	-0.1606	-0.1042	-0.2936	0.7828	1.0000

Adding year dummies and state dummies to control the unobserved effects we have estimated the model but we have found none of the variables significant (results are not shown). To check whether dummy variables are creating multi-collinearity problem or not, we have calculated the variance inflating factor (VIF) for each of the variables and found very high VIF (more than 10.0) for some of the regressors including some state dummy variables (highest VIF 42.67 for *elec*, lowest 1.89 for 2000*dummy* and mean VIF 9.77). So we have dropped the state dummy variables from the model and included regional dummy to capture region specific unobservable effects. To construct the set of regional dummy variables we have divided all the states of India into four regions[8], according to their geographical location

[8]along with these four region, in another exercise, we have considered five regions with Madhya Pradesh and Uttar Pradesh as central region, but the results differ a little and so

1. Northern Region: Haryana, Punjab, Rajasthan, Uttar Pradesh.
2. Southern Region: Andhra Pradesh, Karnataka, Kerala, Tamil Nadu.
3. Eastern Region: Bihar, Orissa, West Bengal, Assam.
4. Western region: Gujarat, Madhya Pradesh, Maharashtra.

After including the regional dummies, we have estimated the model and calculated the VIF. This time we found that the VIF for each of the variables, including year dummies and regional dummies, are less than 7.0, which is shown in Table-3.

Table-3: VIF of the regressors

variables	VIF
nsdip	1.81
profit	1.80
road	1.68
labour	1.68
pcnsdp	3.50
elec	6.41
1984 dummy	1.89
1985 dummy	1.89
1986 dummy	1.89
1987 dummy	1.89
1988 dummy	1.89
1989 dummy	1.89
1990 dummy	1.89

continued on the next page

it is not reported here.

Table-3 (continued)

variables	VIF
1991 dummy	1.89
1992 dummy	1.89
1993 dummy	1.89
1994 dummy	1.89
1995 dummy	1.89
1996 dummy	1.89
1997 dummy	1.89
1998 dummy	1.89
1999 dummy	1.89
2000 dummy	1.89
north region dummy	3.49
south region dummy	2.75
west region dummy	5.03
mean VIF	2.18

But firms' expectation about future values of explanatory variables may depends not only on levels of those variables in recent years but also the magnitude and direction of changes in those variables. To take care of such possibility, we have in an alternative specification taken, in addition, four year average of annual changes in all regressors except *nsdip*. Let us call this model as Model-2[9].

[9]We have dropped *elec* from Model-2 to avoid the loss of degree of freedom.

Model-2

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2(x_{2it} - \bar{x}_{2t}) + \alpha_2(\Delta x_{2it} - \overline{\Delta x_{2t}}\ ) + \beta_3(ln(x_{3it}) - ln(\bar{x}_{3t}))$$
$$+\alpha_3(\Delta ln(x_{3it}) - \overline{\Delta ln(x_{3t})}\ ) + \beta_4(x_{4it} - \bar{x}_{4t}) + \alpha_4(\Delta x_{4it} - \overline{\Delta x_{4t}}\ )$$
$$+\beta_5(ln(x_{5it}) - ln(\bar{x}_{5t})) + \alpha_5(\Delta ln(x_{5it}) - \overline{\Delta ln(x_{5t})}\ ) + c_i + T_t + u_{it}$$

where Δx_{2it} : average of changes in profit share in state i from year $t-4$ to $t-1$.

Δx_{3it} : average of changes in road length per square km in state i from year $t-4$ to $t-1$.

Δx_{4it} : average of changes in percentage of workers involved in labour conflicts in state i from year $t-4$ to $t-1$.

Δx_{5it} : average of changes in per capita net state domestic product in state i from year $t-4$ to $t-1$.

Apart from the most general[10] estimation procedure of panel data model, we have estimated Model-1 and Model-2 using both random effect and fixed effect models. But the Hausman test indicates that we should use the random effect model. So in chapter-4 we have reported results of the random effect model regression only.

[10] i.e. under the weakest set of assumptions.

Chapter-4

Results and Conclusion

In this chapter we have explained the findings of our models in section 4.1 and the summary and conclusions of our study have been presented in section 4.2.

4.1 Results of the study

The results of the OLS regressions of Model-1 and Model-2 are shown in Table-5. The results of the Model-1 are given in column 1 and that of Model-2 are given in column 2.

We have found that the variable *nsdip* has significant positive effect on *inv* at 1% level of significance in both the models. Positive coefficient of *nsdip* implies that if net state domestic industrial product (as a proportion of the total for all the fifteen states) in any state is higher, then investment as a proportion of total investment would *ceteris paribus* be higher in that state. Since *nsdip* indicates the degree of industrialization of a state relative to other states, the positive coefficient implies that the regions with well developed industrial sectors were getting proportionately more investment over the concerned period. In other words we can say that, in context of India, investors prefer to invest in that state which is highly industrialized.

The variable *labour* has significantly negative effects on the distribution of the quantum of investment at 1% level of significance in Model-1 and at 5% level of significance in Model-2. Since the variable *labour* is defined as the percentage of workers involved in conflicts, the negative effects implies that, in any state, the more proportion of workers involved in conflicts, the less investment in that state. But the variable $\Delta labour$, which gives the four year average of annual changes in the percentage of workers involved in labour disputes, in Model-2 is not significant.

The two variables *road* and *elec*, proxying the level of infrastructure, are insignificant in both the models. The sign of the coefficients of both the variables are negative in both the models. But the four year average of yearly changes in length of road per square km has significant positive effect on the distribution of investment at 10% level of significance. The positive effects implies that expectation of better infrastructure, in future, in a state would leads to a greater proportion of investment in that state.

The unexpected result of our study is the negative effects of profit share which is weakly significant at 10% level of significance in Model-1 but insignificant in model-2. The sign of the coefficient of the yearly changes in profit share is positive but insignificant in both the model. The coefficients of the variable *pcnsdp* has expected positive sign but its yearly changes has a negative coefficient.

We have found none of the time dummies and the regional dummies, which are the dummy for the time specific and state specific unobservable effects, to have significant effects except the dummy for the Northern region which has a negative coefficient at 10% level of significance in Model-2.

The results of the estimation of Model-1 and Model-2, with random effect approach, are shown in Table-6. In both the models, only *nsdip* is significant with positive coefficient at 1% level of significance. None of the other variables are significant.

Table-5 : Results of the OLS regression

	model-1	model-2
nsdip	.7180068 [a] (4.10)	.7863141 [a] (4.16)
profit	−.0006294 [c] (−1.86)	−.0003651 (−1.09)
Δ profit		.0001325 (0.07)
road	−.008052 (−0.83)	−.0102338 (−1.26)
Δroad		.3660714[c] (2.01)
labour	−.000564 [a] (−2.96)	−.0005886 [b] (−2.59)
Δ labour		.0003075 (0.76)
pcnsdp	.0090256 (0.28)	.0026261 (0.10)
Δ pcnsdp		−.1366654 (−0.48)
elec	−.0158822 (−0.70)	
1984 dummy	.000279 (−0.00)	
1985 dummy	.0000825 (−0.00)	−.0004445 (−0.01)
1986 dummy	.0000743 (−0.00)	−.0004179 (−0.01)
1987 dummy	−.0000293 (−0.00)	−.0004402 (−0.01)
1988 dummy	−.0000465 (−0.00)	−.000409 (−0.01)
1989 dummy	−.0004458 (−0.01)	−.0005884 (−0.01)
1990 dummy	−.0007716 (−0.01)	−.0007135 (−0.01)
1991 dummy	−.0008125 (−0.01)	−.0007204 (−0.01)

continued on the next page

Table-5 (continued)

variables	model-1	model-2
1992 dummy	−.0008402 (−0.01)	−.0007168 (−0.01)
1993 dummy	−.0005086 (−0.00)	−.0004456 (−0.00)
1994 dummy	−.000412 (−0.01)	−.0004645 (−0.01)
1995 dummy	−.0000896 (−0.00)	−.000277 (−0.01)
1996 dummy	.0002366 (−0.00)	−.000075 (−0.00)
1997 dummy	.0001067 (−0.00)	−.0001665 (−0.00)
1998 dummy	−.0000162 (−0.00)	−.0002028 (−0.00)
1999 dummy	−.0001305 (−0.00)	−.0002143 (−0.00)
2000 dummy	−.0004132 (−0.01)	−.0002707 (−0.00)
northern region dummy	−.0235696 (−1.04)	−.0384827 [c] (−2.11)
southern region dummy	.0177476 (1.03)	.010119 (1.05)
western region dummy	.0380363 (1.55)	.0115453 (0.48)
constant	.0089805 (0.14)	.0180187 (0.35)

Note:The dependent variable is *inv* in all the two models.

t statistics are given in parenthesis

a: significant at 1% level of significance.

b: significant at 5% level of significance.

c: significant at 10% level of significance.

Table-6 : Results of the random effect model regression

	model-1	model-2
nsdip	.7007877 [b] (3.39)	.7578643 [b] (3.17)
profit	−.0006361 (−1.30)	−.0003274 (−0.63)
Δprofit		.0000344 (0.02)
road	−.0081849 (−0.83)	−.0123315 (1.03)
Δroad		.3100172 (0.90)
labour	−.0003885 (−1.29)	−.0002844 (−0.62)
Δlabour		.0000896 ((0.17))
pcnsdp	.0199796 (0.55)	.0080206 (0.18)
Δpcnsdp		−.1352923 (−0.32)
elec	−.0245049 (−1.20)	
1984 dummy	.0004782 (0.01)	
1985 dummy	.0002932 (0.01)	−.0005489 (−0.01)
1986 dummy	.0002396 (0.00)	−.0005366 (−0.01)
1987 dummy	.0000586 (0.00)	−.0005842 (−0.01)
1988 dummy	$-8.83e^{-06}$ (−0.00)	−.0005641 (−0.01)
1989 dummy	−.0004733 (−0.01)	−.0007334 (−0.01)
1990 dummy	−.0008618 (−0.02)	−.0008459 (−0.02)
1991 dummy	−.0009104 (−0.02)	−.000847 (−0.02)

continued on the next page

Table-6 (continued)

variables	model-1	model-2
1992 dummy	−.0009471 (−0.02)	−.000838 (−0.02)
1993 dummy	−.0006224 (−0.01)	−.0005555 (−0.01)
1994 dummy	−.0004027 (−0.01)	−.0005382 (−0.01)
1995 dummy	$7.29e^{-06}$ (0.00)	−.0003168 (−0.01)
1996 dummy	.0004025 (0.01)	−.0000869 (−0.00)
1997 dummy	.0003221 (0.01)	−.0001452 (−0.00)
1998 dummy	.0001719 (0.00)	−.0001736 (−0.00)
1999 dummy	$1.84e^{-06}$ (0.00)	−.0001866 (−0.00)
2000 dummy	−.0003739 (−0.01)	−.000232 (−0.00)
northern region dummy	−.0162914 (−0.71)	−.0355779 (−1.19)
southern region dummy	.0225907 (1.34)	.0096378 (0.48)
western region dummy	.0478249 (1.74)	.0154683 (0.46)
constant	.0037487 (0.06)	.0183451 (0.34)

Note:The dependent variable is *inv* in all the two models.

z statistics are given in parenthesis

b: significant at 1% level of significance.

4.2 Summary and Conclusion

Despite almost five decades of continuous efforts on the part of Indian planners and the central government, regional disparity has been a persistent and increasing problem in India. The states which were rich, in terms of PCSDP, were growing at a higher rate than their poor counterpart. Rao et al (1999) found that investment was the most important determinant of growth rate since the mid-1970s. It was also found that, in the 1980s, the agricultural sector was the main source of disparity, followed by the manufacturing sector. But in the 1990s, manufacturing was the main source. We have found that investment disparity across states has been increasing over time since the 1980s.Therefore, in this dissertation our objective has been to find out the factors that affects the regional distribution of investment across the states of India.

Investment theory indicates various factors that affect the distribution of the quantum of investment. According to the neoclassical theory of investment, the distribution of investment depends on the stock of capital goods and the real rate of interest. Because of diminishing returns of capital, the marginal product of capital would be relatively low in a state with a relatively high level of capital stock which may deter investment in that state. So the size of the capital stock affects investment negatively. Since an increase in the real interest rate increases the cost of capital goods, the theory also implies that a region with a greater real rate of interest would have a lower volume of investment. But if interest rates are equalized across regions then there would be no effect on the distribution of investment. Kalecki's theory highlights the importance of the level of savings of firms and changes in business profits as determinants of the volume of investment. High level of savings would reduce the problem of the limited availability of finance, and firms would be able to use internal savings for investment. The savings of firms are closely related to their level of output and profit margin. If the levels of output and /or profit margin increase, firms' profits would increase and so would savings. So regions with higher levels of output and higher levels of profit share will have greater investment. Since increases in profit share increase the expectation of future profitability, regions with greater increases in profit share may be expected *ceteris paribus* have greater quantity of investment. According to Tobin's q-theory of investment, investment depends on the market valuation of capital goods. If the market valuation of a stock of capital goods is greater in a region than in others with the same stock of capital goods, then the former should get a greater quantity of investment.

The literature on the location choice problem indicates various factors that affect the location choice decisions of firms. Among them the effects of higher wages and greater unionisation of labour has been generally found to be negative. The variable per capita income which is taken as a proxy for market situation, has generally shown a positive effect. The effect of variables measuring the level of infrastructure has usually been found positive, except in the study by Sanyal and Menon (2005) in the context of India where the effect was found to be negative. The effects of the existing level of the stock of capital goods has generally been found to be positive.

Our study covers fifteen states of India over the period 1984 to 2001 which is the period of transition from the pre-reform era to post reform era. The fifteen states we have studied cover almost 90 percent of the Indian population and account for almost 90 percent output in the organized industrial sector. These were, more or less, politically stable during the period considered and were not union territory since 1984. We have found that the size of the existing capital goods has significant positive effect on investment. Our results supports the contention of Rao et all (1999) that neoclassical assumption of 'diminishing marginal returns' of capital goods can not be applied in the case of the developing countries like India. Alternatively, it can be said that the benefits from being located in a highly industrialized region outweighs the costs of having a high level of capital stock. The effects of the percentage of workers involved in labour conflicts was found to be negative. If we assume that labour conflict is positively related to the degree of organization of labour, then this lends support to the finding of Sanyal and Menon (2005) who obtained a negative coefficient for the variable - number of unionized workers. Like Sanyal and Menon (2005), we have obtained a negative coefficient for the proxy for infrastructure (although insignificant) but we have found significant positive effects of the variable, change in road length per square km. The surprising result of our study is the significant negative effect of profit share on the distribution of investment.

There are some variables, like area of state, wage rate, government incentives and taxation to the industry, human capital, which we would have liked to include in our study, but could not due to various constraint. Since state areas are fixed, inclusion of this variables would not be appropriate in a panel data model. They are partly accounted for by the regional dummies. The same is true for human capital which can be proxied by the literacy rate, but not available for each of the years of the period considered. State wise average yearly wage rate was not available to us. The problem with government incentives and taxation is its dimension. In India different state governments provide different types of tax concession and incentives to the industry under different heads. So it is difficult to compose the extent of incentives and concessions provided by various states using any single variable or index. There, therefore, remains scope for extending this study in the future to take into account the effects of these additional variables.

References

Barro, Robert J. and Xavier Sala-i- Martin (1992): Convergence, *Journal of Political Economy*, vol. 100, no. 2.

Bartik, Timothy J. (1985): Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States, *Journal of Business and Economic Statistics*, vol. 3, no. 1.

Central Electricity Authority: *All India Electricity Statistics*, various issue, Govt. of India.

Chakravorty, Sanjoy (2003): Industrial Location in Post-reform India: Patterns of Inter-regional Divergence and Intra-regional Convergence, *Journal of Development Studies*,40: 2.

Cheng, Leonard k. and Yum K. Kwan (2000): What are the Determinants of the Location of Foreign Direct Investment? The Chinese Experience, *Journal of International Economics*, 51.

Coughlin, Cletus C., Joseph V. Terja and Vachira Arromdee (1991): State Characteristics and the Location of Foreign Direct Investment Within the United State, *The Review of Economics and Statistics*, 73.

Dasgupta, Dipankar et all (2002): Growth and Interstate Disparities in India, *Economic and Political Weekly*, July 1.

Disdier, Anne-Celia and Thierry Mayer (2004): How Different is Eastern Europe? Structure and Determinants of Location Choices by French Firms in Eastern and Western Europe, *Journal of Comparative Economics* , 32.

Ghosh, Budhadeb et all (1998): Economic Growth and Regional Divergence in India,1960 to 1995, *Economic and Political Weekly*, July 27.

Herrin, Alejandro N. and Ernesto M. Pernia (1987): Factors Influencing the Choice of Location: Local and Foreign Firms in the Philippines, *Regional Studies*, vol. 21.6.

Labour Bureau: *Indian Labour Statistics*, various issue, Ministry of Labour, Govt. of India.

Labour Bureau: *Indian Labour Year Book*, various issue, Ministry of Labour, Govt. of India.

Labour Bureau: *Industrial Disputes in India*, various issue, Ministry of Labour, Govt. of India.

Kalecki, Michal (1971): *Selected Essays on the Dynamics of the Capitalist Economy*, Cambridge University Press.

Mankiw, N. Gregory (2006): *Macroeconomics*, 5th ed, Worth Publisher.

Rao, M Govinda et al(1999) : Convergence of Income Across Indian States: A Divergent View, *Economic and Political Weekly*, March 27.

Romer, D (1986): *Advanced Macroeconomics*, The Macgraw-hill Companies,Inc.

Sanyal, Paroma and Nidhiya Menon (2005): Labor Disputes and the Economics of Firm Geography: A Study of Domestic Investment in India, *Economic Development and Cultural Change*, 53.

Sun, Qian et al (2002) : Determinants of foreign direct investment across

China, *Journal of International Money and Finance*, 21.

Woodward, Douglas p. (1992): Locational Determinants of Japanese Manufacturing Start-ups in the United States, *Southern Economic Journal*,vol. 58, no. 3.