JOURNAL OF DEVELOPMENT ECONOMICS AND FINANCE
Vol. 1, No. 2, 2020, 209-225

URL : www.arfjournals.com

Financial Inclusion in the States of India: A Panel Data Analysis of Accounts Penetration

T. Lakshmanasamy*

** Formerly Professor, Department of Econometrics, University of Madras, Chennai, E-mail: tlsamy@yahoo.co.in*

Article History

Received : 13 July 2020
Revised : 17 July 2020
Accepted : 24 August 2020
Published : 30 December 2020

Key words

Financial inclusion, deposit and credit penetration, panel estimation.

Abstract: The level of financial inclusion in India is not only low but also differ widely across states of India. This paper aims to identify and estimate the effects of the determinants of deposit and credit penetration in states of India, the indicators of financial inclusion, using state-wise panel data over the period 2001-2013, applying panel regression methods. The paper observes a relatively high level of bank accounts holding in states like Himachal Pradesh and a low level of financial activities in states like Nagaland and Bihar. The panel estimates of accounts penetrations reveal a significant effect of NSDP per capita income on deposit penetration, while beyond the income level, bank branch networking and access to banks also determine the level of credit penetration in states of India. The intensity of financial inclusion in Indian states depends on not only banking sector variables but also other state-level development and economic factors. This study reveals that in deposit penetration, income and industrialisation of states play a vital role and apart from income and industrialisation, population density and bank branch networking in states also matter for credit penetration.

INTRODUCTION

A strong and efficient financial system is required not only in developing countries but also in developed countries to achieve economic growth and development. A well-functioning financial system is equally important for financial inclusion, which is complementary to inclusive growth in developing countries. Financial inclusion is vital in improving financial stability also. Financial inclusion is about (i) broadening of financial services to those people who do not have access to financial services, (ii) deepening of financial services for people who have minimal financial services, and (iii) greater financial literacy and consumer protection. Any inefficiency in these areas leads to financial exclusion of the poor and weaker sections of the population. It is estimated that globally over 2.5 billion people are excluded from access to financial services of which one third is in India. The Eleventh Five Year Plan (2007-12) of India envisions inclusive growth as a key policy objective.

The Reserve Bank of India views financial inclusion as the process of ensuring access to appropriate financial products and services needed by vulnerable groups such as weaker sections and low-income groups at an affordable cost fairly and transparent manner by mainstream institutions. In India, the focus of the financial inclusion is confined to ensuring bare minimum access to a savings bank account to all. Internationally, financial inclusion has been viewed from a much wider perspective. Simply having a current/savings account on its own is not regarded as an accurate indicator of financial inclusion. Financial inclusion should offer access to a range of financial services including savings, long and short-term credit, insurance, pensions, mortgages, money transfers, etc. importantly at an affordable cost. The commonly recognised six pillars strategy for financial inclusion is (i) universal electronic bank account, (ii) access to banking payment services, (iii) access to credit/loan, (iv) access to investment, (v) access to insurance, and (vi) consumer protection.

In a vast and federal country like India, the determinants of financial inclusion at the aggregate level are complex and vary by states, as states in India have differing socioeconomic, demographic and cultural patterns and also their ways and means for financing the poor and vulnerable sections of its population. Therefore, the problem of financial inclusion, as well as exclusion varies between states. Table 1 and Figures 1 and 2 present the extent of financial inclusion in terms of deposit and credit penetration in states of India. It can be observed that both deposit and credit penetrations have similar patterns, Himachal Pradesh having the highest penetration while Nagaland having lowest penetration.

Given such wide cross states differences in accounts penetration, it is important to examine the determinants of financial inclusion in states, not merely at the individual or national levels. Therefore, this study attempts to analyse the determinants of financial inclusion in states of India. The main objectives of the study are to identify the determinants of financial inclusion in states of India and to understand the intensity of financial inclusion across states in India. This study uses panel data for 26 states for 13 years from 2001 to 2013. Empirically this study uses a multidimensional approach to measure financial intensity using accounts penetration, by deposit and credit penetration. In the empirical estimation, panel data regression techniques like fixed and random effects models are used.

REVIEW OF LITERATURE

Sarma (2008; 2015) devises distance-based multidimensional index of financial inclusion (IFI) for an inclusive financial system based on banking penetration, availability of banking services and usage, using data from the World Development Indicators of the World Bank and International Financial Statistics of the IMF. The financial (banking) penetration dimension is constructed from bank deposit accounts including checking (or current), savings and time deposit accounts for business,

Table 1: Mean Level of Financial Inclusion in the States of India

State/UT	*Deposit penetration*	*Credit penetration*	*State*	*Deposit penetration*	*Credit penetration*
A&N Islands	527.19	45.33	Maharashtra	614.01	141.55
AP	630.35	133.08	Manipur	748.05	106.75
Assam	2514.14	271.12	Meghalaya	350.74	43.80
Bihar	272.91	35.31	Nagaland	197.64	30.51
Delhi	1480.43	126.80	Orissa	403.87	73.93
Goa	1924.75	110.52	Pondicherry	844.28	144.21
Gujarat	582.06	50.82	Punjab	2457.54	195.58
Haryana	648.53	61.78	Rajasthan	1556.36	189.21
HP	7404.63	669.58	Sikkim	453.66	53.91
J&K	2307.79	181.75	Tamil Nadu	667.54	193.69
Karnataka	694.32	128.22	Tripura	429.25	85.73
Kerala	828.99	158.94	Uttar Pradesh	448.65	43.93
MP	360.40	44.45	West Bengal	514.33	45.95

Figure 1: Mean Deposit Penetration in the States in India

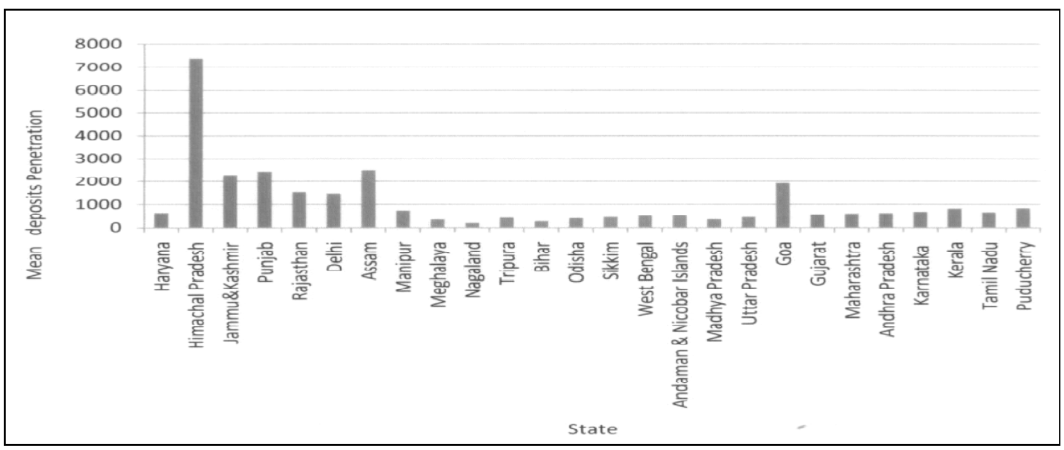
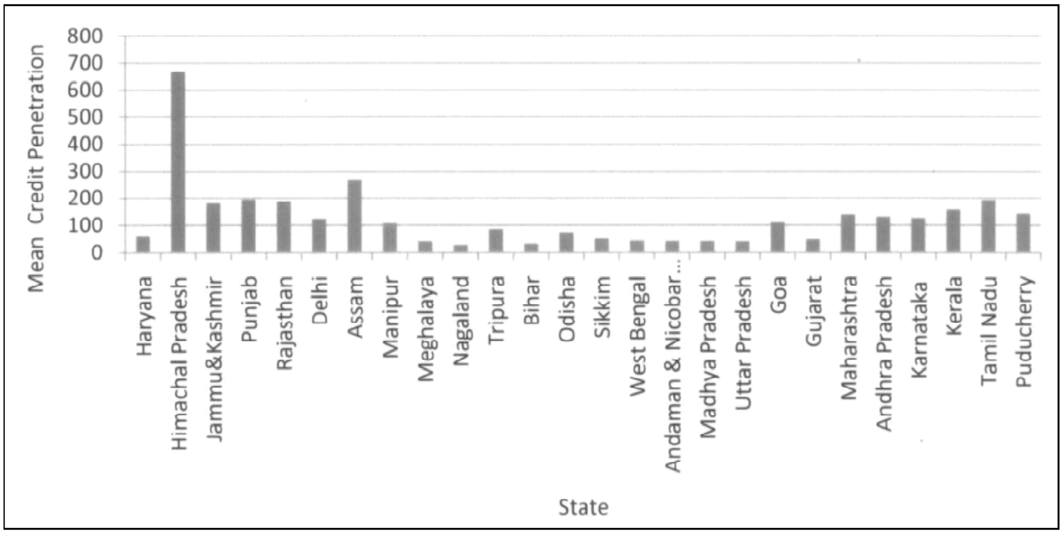


Figure 2: Mean Credit Penetration in the States of India



individuals and others. The availability of banking services dimension is constructed from deposit money in commercial banks and other financial institutions that accept transferable deposits, such as demand deposits. The usage of banking services dimension is constructed from domestic claims on the private and residential sector and total deposits. The constructed IFI reveal that Spain leads with the highest value of IFI and India is ranked at 31 among 55 countries.

Sarma and Pais (2011) present a cross country empirical analysis of the relationship between financial inclusion and development using the index of financial inclusion proposed by Sarma (2008). The analysis shows that income is positively associated with the level of financial inclusion, and inequality, literacy and urbanisation also influence financial inclusion. Further, physical infrastructure for connectivity and information are important for financial inclusion. Among the banking sector variables, NPA and CAR are negatively associated while the interest rate is not associated with financial inclusion. While government ownership of banks is not important for financial inclusion, foreign ownership is negatively associated. These results show that financial exclusion is a refiection of social exclusion, as countries having low GDP per capita, relatively higher levels of income inequality, low rates of literacy, low urbanisation and poor connectivity are financially less inclusive.

Kumar (2013) attempts to understand the behaviour and determinants of financial inclusion in India using state-level panel data spanning over I995 to 2008.

The panel fixed and random effects estimates of deposit penetration and credit penetration show a negative infiuence of population density on deposit penetration, contrary to usual perception, but a significant positive infiuence on credit penetration. The level of per capita NSDP, level of industrialisation and employment base are significantly related to deposit penetration. The average population per branch has a negative infiuence on credit penetration. The time effects are significant in explaining the deposit and credit penetration indices.

Chakravarty and Pal (2013) analyse financial inclusion in across countries and states of India using the IFI constructed on geographic branch penetration, demographic branch penetration, geographic ATM penetration, deposit accounts per capita, credit accounts per capita, deposit income ratio and credit income ratio. The principal component analysis shows wide variations in financial inclusion across countries and states. The financial inclusion across states in India show that Delhi and consistently maintain their ranks over time. The contribution of geographical penetration of bank branches to overall achievement is least.

Chithra and Selvam (2013) analyse the determinants of financial inclusion through inter-state variations in access to finance, using the composite financial inclusion index (IFI) developed by Sarma (2008). Using index regression analysis with a set of socio-economic, infrastructure and banking variables, the study finds that GDP per capita, literacy rate, internet, phone facilities, rural population, road network and deposit penetration have a significant positive association with IFI, while credit penetration has a negative relation with IFI. The results indicate wide inter-state variation in the level of financial inclusion, and connectivity and information play an important role in financial inclusion in India.

Gupta *et al.* (2014) analyse the extent of financial inclusion across 28 states and 6 regions of India computing IFI using the three dimensions of financial inclusion viz. penetration, availability and usage of banking services. Based on IFI, the paper concludes that the states of Goa, Punjab and Kerala are the most financially inclusive states of India. It is also observed that although various measures have been implemented to increase financial inclusion, a large population of India does not have access to formal financial system.

Nandru *et al.* (2016) attempt to identify the determinants of financial inclusion by providing evidential support for south Indian states by simple regression on an index of financial inclusion (IFI) derived from Crisil's Inclusix. The variables considered are population size, gender ratio, branch penetration, deposit to credit penetration ratio and literacy ratio. The estimated regression results indicate that population size, gender ratio, branch penetration, and deposit to credit penetration ratio are significant determinants of financial inclusion in states of India.

Inoue and Harnori (2012) examine whether financial inclusion contributes to poverty reduction in India, using unbalanced panel data for 28 states and union territories between 1973 and 2004. on to find whether financial deepening and economic growth alleviate poverty. The study uses poverty ratio as the dependent variable and explanatory variables are the measures of financial deepening, ratio of credit to GDP and ratio of deposits to GDP. The dynamic generalised method of moments estimates shows that financial deepening and economic growth improve the poverty ratio whereas international openness and infiation worsen poverty ratio.

Dixit and Ghosh (2013) attempt to understand the role of financial inclusion as an instrument to attain inclusive growth in Indian states. The dendrogram of the average linkage between natural hierarchical clusters with parameters like GDP per capita, literacy rate, unemployment rate and index of financial inclusion is considered. The dendrogram analysis shows that at five rescaled distance three major clusters emerge from considered parameters. The states with high GDP per capita income and literacy rate have high financial inclusion. The states with low GDP per capita income account for low financial inclusion.

Empirical results similar to India are observed in other developing countries also. Marin and Schwabe (2013) observe a positive relationship between bank competition and penetration of bank accounts at the municipal level in Mexico. Using a two stage estimation of account penetration, the study finds that in markets in which the provision of bank services is more concentrated, people are less likely to use financial services.

Camara *et al.* (2014) examine financial inclusion in Peru based on micro-data of Global Findex for Peru. The probit estimates on whether a household has an account or not show that traditional factors such as being a woman, living in a rural area or having a low income and educational level may reduce the likelihood of being included in the formal financial system. Education is more important for enterprises than for households in fostering financial inclusion. More than 50 percent of the unbanked perceive the lack of money and high cost of financial services as the main obstacles for the financial inclusion.

Tuesta *et al.* (2015) analyse the three dimensions viz. access, use and barriers determining financial inclusion in Argentina using Global Findex (2012), a financial survey by the World Bank, in an attempt to understand why individuals do not participate in formal financial system. The probit estimates of whether a household or an enterprise has a bank account or not with banking products and services reveal that level of education, income, and age are important variables. Although income level may be a structural problem, age might reflect the absence of financial products to the meet the needs of different groups.

Alter and Yontcheva (2015) study whether structural characteristics of regions hamper financial inclusion and development, measured by the ratio of private credit to GDP and financial development gap, comparing Central African Economic and Monetary Community (CEMAC) with its peers from Sub-Saharan Africa (SSA). The panel fixed effects estimates identify that macroeconomic variables like inflation, income, new technology, bank operational costs explain the most of the private credit to GDP ratio in CEMAC. Financial development is positively linked to the number of bank branches, availability of credit information, registry coverage and negatively impacted by bank operational costs, cost-income ratio and poverty headcount.

DATA AND METHODOLOGY

This study uses state-wise panel data for 26 states over 13 years from 2001 to 2013 collected from various sources, consisting of an unbalanced panel of 290 observations. This study uses deposit penetration and credit penetration as dependent variables and the independent variables are population density, the average population per branch (APPB), per capita NSDP, credit-deposit ratio, number of factories and number of employees. The deposit penetration (DP) is the number of deposit accounts per thousand population. Deposit penetration indicates the accessibility of basic banking services, like having a bank account, loan, etc. The credit penetration (CP) is the number of loan/credit accounts per thousand population. Credit penetration indicates the availability of loans and the volume of credit circulated in the economy. The CRISIL Inclusix (2013) which measures the extent of financial inclusion at geographical level takes deposit, credit and branch penetration as important indicators of financial inclusion. The independent variables of the study include population density, per capita NSDP which is a proxy for income, credit-deposit ratio, the average population per branch (APPB), number of factories as an indicator of industrialisation and industrial employment base.

The states included are Haryana, Himachal Pradesh, Jammu & Kashmir, Punjab, Rajasthan, Delhi, Assam, Manipur, Meghalaya, Nagaland, Tripura, Bihar, Orissa, Sikkim, West Bengal, Andaman & Nicobar Islands, Madhya Pradesh, Uttar Pradesh, Goa, Gujarat, Maharashtra, Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, and Pondicherry. The data on the number of deposit accounts, credit accounts, number of branches, amount of deposits and loans are obtained from the Basic Statistical Returns (BSR) of the Scheduled Commercial Banks in India published by the Reserve Bank of India. The data on per capita Net State Domestic Product (NSDPpc) is collected from the RBI Database on Indian Economy, RBI 1980-81 base year. As population census is done once in ten years, projected population data from the Population Projections for India and States is used. The variables number of industries and the number of employees are from the Annual Survey of Industries.

PANEL DATA METHOD

Panel data consists of repeated observations over time, a repeated cross-section of time series. If some observations for some variables are missing during the period the data set is an unbalanced panel data. The panel data allows great fiexibility in modelling differences in behaviour across individuals and therefore allows studying individual dynamics. By controlling individual unobserved heterogeneity, the panel estimates are more efficient.

The basic panel regression model can be specified as:

$$y_{it} = \alpha z_i + \beta x_{it} + \lambda_i + u_{it} \; i = 1,\dots,n \; t = 1,\dots, T_i \tag{1}$$

where z includes observed time-invariant individual or group-specific variables such as race, sex, location, etc. along with the constant term. The λ_i are the unobserved individual or group-specific effects such as family-specific characteristics, individual heterogeneity in skill preferences, etc. If all the observed and unobserved individual or group-specific heterogeneity are constant over time, the model is a classical regression model, then the entire model can be estimated by the ordinary least squares method. The complications arise when the unobserved heterogeneity is correlated with the error term, then OLS estimation will produce biased (inconsistent and inefficient) estimates. Generally, a panel data is estimated by pooled data, least squares dummy variables, fixed effects and random effects regression models.

Pooled Regression: If z_i and λ_i is constant and uncorrelated with x_i, then OLS estimation provides consistent and efficient estimates of common α and the slope vector β. The assumptions of the pooled model are:

$$y_{it} = \alpha + \beta x_{it} + u_{it} \tag{2}$$

$$E(u_{it} \mid x_{it}) = 0 \quad Var(u_{it} \mid x_{it}) = E(u_{it}^2) = \sigma_u^2 \quad Cov(u_{it}, x_{it}) = 0 \tag{3}$$

The pooled regression is also called a population-averaged model as the presence of any latent heterogeneity is averaged out. To the pooled data, the least squares regression is applied under the assumptions of zero conditional mean of the error, homoscedasticity, independence across observations and strict homogeneity of x.

Fixed Effects Regression: If λ_i is correlated with x_i, then the least squares estimator of fi is biased and inconsistent as a consequence of an omitted variable bias. In this instance, the observable individual effects are assumed to be fixed or remain constant over time and such a fixed effects model can be specified as:

$$y_{it} = \alpha_i + \beta x_{it} + u_{it} \tag{4}$$

where $\alpha_i = (\alpha + \lambda_i)$ is a group-specific constant term. However, the omitted unobservable individual effect λ_i may be correlated with x_i. That is:

$$E(\lambda_i \mid x_{it}) = g(x_i) \Rightarrow E(u_{it} \mid x_{it}) \neq 0 \tag{5}$$

As the conditional mean is the same in every period, the model can be written as:

$$y_{it} = \beta x_{it} + g(x_i) + [\lambda_i - g(x_i)] + u_{it} \tag{6}$$

$$y_{it} = \alpha_i + \beta x_{it} + [\lambda_i - g(x_i)] + u_{it} \tag{7}$$

By construction the bracketed term is uncorrelated with x and is therefore absorbed in the disturbance term, giving the fixed effects panel regression model. The fixed effects regression model captures the individual differences in the constant term α_i and each α_i is treated as an anonymous parameter to be estimated along with the slope vector β. The fixed effects regression models the differences between cross-sectional units or groups strictly as parametric shifts of the regression function i.e. λ_i is non-stochastic and not that any variable is fixed.

The fixed effects model can be estimated in three ways: least squares dummy variable, within-group and between-group regressions. The least squares dummy variable (LSDV) regression specifies a set of dummy variables for each cross-sectional unit:

$$y_{it} = \beta x_{it} + d\lambda_i + u_{it} \tag{8}$$

The dummy coefficients are shifts in intercepts only leaving slope parameters unaltered. The fixed effects within-group regression uses group-means deviations:

$$(y_{it} - \bar{y}_i) = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_i) \tag{9}$$

The slope parameters are estimated using the within-group averages as:

$$\hat{\beta}_W = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it}-\bar{x}_i)(y_{it}-\bar{y}_i)}{\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it}-\bar{x}_i)^2} \tag{10}$$

The fixed effects between-group regression uses variations of the group-means around overall mean:

$$(\bar{y}_i - \bar{y}) = \beta(\bar{x}_i - \bar{x}) + (\bar{u}_i - \bar{u}) \tag{11}$$

The parameters of the fixed effects between group panel regression is estimated as:

$$\hat{\beta}_B = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}(\bar{x}_i-\bar{x})(\bar{y}_i-\bar{y})}{\sum_{i=1}^{n}\sum_{t=1}^{T}\bar{x}_i-\bar{x})^2} \tag{12}$$

But, estimating so many constant terms as there are cross-sectional observations is costly in terms of degrees of freedom lost.

Random Effects Regression: If the unobserved individual heterogeneity, however formulated, is assumed to be uncorrelated with x, then the individual specific constant terms λ_i may be treated as randomly distributed across cross-sectional units in the same way as y_{it} and x_{it}. As there is no need to estimate each of the i separately, the number of parameters to be estimated are drastically reduced. Thus, a linear regression model can be estimated with a composite disturbance term that may be consistent, although inefficiently, by least squares. The random-effects or error components regression model can be specified as:

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it} \tag{13}$$

where $\varepsilon_{it} = (\lambda_i + u_{it})$ and the single constant term is the mean of the unobserved heterogeneity. The component λ_i is the random heterogeneity specific to the i^{th} observation and is constant through time. The assumptions of the random effects regression model are:

$$\begin{aligned}
&E(\varepsilon_{it} \mid x_{it}) = E(\lambda_i \mid x_{it}) = 0 \\
&E\left(\varepsilon_{it}^2\right) = \sigma_u^2 + \sigma_\lambda^2 \\
&E(\varepsilon_{it}\varepsilon_{is} \mid x_{it}) = \sigma_\lambda^2 \quad \textit{if } \ t \neq s \\
&E(\varepsilon_{it}\varepsilon_{is} \mid x_{it}) = \sigma_\lambda^2 \quad \textit{for al t and s if } \ i \neq j
\end{aligned} \tag{14}$$

The variance-covariance matrix of cross-sectional unit i is specified as:

$$E(\varepsilon_i \varepsilon_i' \mid x_{it}) = \Sigma = \begin{bmatrix} \sigma_u^2 + \sigma_\lambda^2 & \sigma_\lambda^2 & \dots & \sigma_\lambda^2 \\ \sigma_\lambda^2 & \sigma_u^2 + \sigma_\lambda^2 & \dots & \sigma_\lambda^2 \\ \dots & \dots & \dots & \dots \\ \sigma_\lambda^2 & \dots & \dots & \sigma_u^2 + \sigma_\lambda^2 \end{bmatrix} \tag{15}$$

As cross-section units *i* and *j* are independent, the covariance matrix of the disturbance term for the full nT observations is:

$$\Omega = \begin{bmatrix} \Sigma & 0 & \dots & 0 \\ 0 & \Sigma & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \Sigma \end{bmatrix} = I_n \otimes \Sigma \tag{16}$$

Then, the estimates of slope parameters $\hat{\beta}$ are obtained from the generalised least squares (GLS) estimation as:

$$\hat{\beta} = (x'\Omega^{-1}x)'x'\Omega^{-1}y = [\textstyle\sum_{i=1}^{n} x_i'\Omega^{-1}x_i]^{-1} \sum_{i=1}^{n} x_i'\Omega^{-1}y_i \tag{17}$$

Thus, the GLS estimator, like the OLS estimator, is a matrix weighted average of within and between cross-sectional units estimators.

Hausman Specification Test

From a purely practical standpoint, which panel regression is to be used in estimation. As the random effects model has the advantage of incorporating the randomness of the individual heterogeneity, the random effects model gives consistent, albeit inefficient, estimates of the parameters. However, if there are omitted variables, the individual specific component λ_i might be correlated with the independent variables in the random effects model, to which the fixed effect model is robust. At the same time, the fixed effects model estimates are consistent, but inefficient, as the λ_i are assumed to be constant and uncorrelated with the error term. Therefore, the choice between fixed effects and random effects panel regression estimation methods is critical. Hausman (1978) has devised a test to choose between the two on the basis that under the hypothesis of no correlation [$Cov(\lambda_i, x_{it}) = 0$], the OLS, fixed effects and random effects estimators are consistent, but OLS is inefficient, whereas under the alternative, the fixed effect is consistent, but the random effect is not. Therefore, under the null hypothesis, the two estimates should not differ systematically and hence a test is proposed on the difference.

The Hausman specification test tests the covariance matrix of the difference vector$[\hat{\beta}_{FE} - \hat{\beta}_{RE}]$ for orthogonality of the common effects and the regressors:

$$Var[\hat{\beta}_{FE} - \hat{\beta}_{RE}] = var(\hat{\beta}_{FE}) + var(\hat{\beta}_{RE}) - cov(\hat{\beta}_{FE}, \hat{\beta}_{RE}) - cov(\hat{\beta}_{RE}, \hat{\beta}_{FE}) \quad (18)$$

If there is no difference between the two estimators, as Hausman shows, then the covariance of an efficient estimator with its difference from an inefficient estimator is zero:

$$Cov[(\hat{\beta}_{FE} - \hat{\beta}_{RE}), \hat{\beta}_{RE}] = \mathrm{cov}(\hat{\beta}_{FE} - \hat{\beta}_{RE}) - \mathrm{var}(\hat{\beta}_{RE}) = 0 \quad (19)$$

$$Cov(\hat{\beta}_{FE} - \hat{\beta}_{RE}) = \mathrm{var}(\hat{\beta}_{RE}) \quad (20)$$

$$ar[\hat{\beta}_{FE} - \hat{\beta}_{RE}] = \mathrm{var}(\hat{\beta}_{FE}) - \mathrm{var}(\hat{\beta}_{RE}) = \boldsymbol{\Omega} \quad (21)$$

where is the covariance matrix for the test. Under the null hypothesis, the chi-square test is based on the Wald criterion:

$$H = [\hat{\beta}_{FE} - \hat{\beta}_{RE}]' \hat{\Omega}^{-1} [\hat{\beta}_{FE} - \hat{\beta}_{RE}] \sim \chi_k^2 \quad (22)$$

Under the null hypothesis, both the estimators are unbiased and consistent. But the random effects estimator is more efficient and the standard error $(\hat{\beta}_{FE})$>standard error$(\hat{\beta}_{RE})$. The null hypothesis and alternative hypothesis for the Hausman test are specified as $H_0 : \hat{\beta}_{FE} = \hat{\beta}_{RE}$ *and* $H_1 : \hat{\beta}_{FE} \neq \hat{\beta}_{RE}$. When the computed value of the testing statistic is greater than the critical fifi2 value, the null hypothesis of the random effects model is rejected and the preferred specification for the data is the fixed effects model.

Empirical Analysis

Table 2 presents the descriptive statistics of the variables used in the study. In the unbalanced panel of 290 observations, the mean of aggregate deposit penetration is 1148.54 and the mean of aggregate credit penetration is 129.48. The average population density is 810 people per square kilometer. Per capita NSDP is Rs. 5494.14. The mean ratio of credit accounts to deposit accounts is 0.52. The mean 11.42 of the average population per branch indicates that at least 11000 population must be covered by a single branch, a large number.

Variable *Definition* *Mean* *Std. dev.*

Table 2: Descriptive Statistics of Variables

Variable	*Definition*	*Mean*	*Std. dev.*
DP	Number of deposit accounts per 1000 population, a measure of the accessibility of banking services	1148.54	1468.34
CP	Number of credit accounts per 1000 population, a measure of the availability of loans and volume of credit	129.48	130.46
NSDPpc	State income per capita, a measure of financial ability of states	5494.14	3410.67
CDR	Ratio of amounts of credit to the amount of deposits, a measure of the flow of money in the economy and accessibility and availability of banking sector to sections of the society	0.52	0.32
APPB	Average number of people per bank branch, a measure of geographical penetration of branches and reach of banking services	11.42	6.51
PD	Population per square kilometer	0.81	2.07
NF	Number of factories, a measure of industrialisation	6283.76	7538.20
EMP	Number of industrial employees, a measure of financial ability of people	479746.98	922187.59
N	Number of observations	338	

contd. table 2

The estimating empirical equations of deposit and credit penetration are specified as:

$$DP_{it} = \alpha_0 + \alpha_1 NSDP_{it} + \alpha_2 CDR_{it} + \alpha_3 APPB_{it} + \alpha_4 NF_{it} + \alpha_5 PD_{it} + \alpha_6 EMP_{it} + \varepsilon_{it} \quad (23)$$

$$CP_{it} = \beta_0 + \beta_1 NSDP_{it} + \beta_2 CDR_{it} + \beta_3 APPB_{it} + \beta_4 NF_{it} + \beta_5 PD_{it} + \beta_6 EMP_{it} + \varepsilon_{it} \quad (24)$$

Table 3 presents the panel regression results of deposit penetration in states of India. In the estimated pooled results NSDP per capita and APPB significantly impact negatively deposit penetration. However, due to pooled dataset, observations for a state may not be independent and there may also be time effects. Hence, the panel fixed effects and random effects regressions are used to control the state-wise heterogeneity due to differences in economic, social and demographic factors across regions over time. In the fixed effects regression, the variables per capita NSDP and number of factories show significant positive impact on deposit penetration. An increase in per capita NSDP increases bank deposits almost by 9 percent. In the random effects estimates, NSDP per capita and APPB are the significant influence deposit penetration. While the effect of NSDP per capita on deposit penetration is positive, the impact of APPB is negative. The negative effect of average population per branch points out in financial inclusion parlance that the poor or dismal penetration of branches, depriving people of access to formal sources of finance.

Generally, population and industrialisation are expected to increase the need for banking services. Contrary to this, the effect of population density and industrial employment on deposit penetration are negative but statistically insignificant. Further, the credit-deposit ratio is also an insignificant determinant of deposit penetration. The rho value, the inter-class correlation of the panel model, indicate that more than 94 percent of the variance in deposit penetration in states is due to differences across panels. Thus, there exists a wide inequality in deposit penetration among states in India. While in Himachal Pradesh has on average 7405 bank accounts per thousand population, almost polar opposites. The within and between R-square values show that around 23 percent of variations in deposit penetration is explained by the differences in factors within states and another 20 percent is due to variations in factors across the states in India.

To identify the appropriateness of fixed effects versus random effects regression, Hausman's specification test is used. The null hypothesis of Hausman's test is that

Table 3: Panel Pooled OLS, Fixed Effects and Random Effects Regression Estimates of Deposit Penetration in the States of India

Dependent variable: Deposit penetration

Variable	*Pooled OLS*	*Fixed effects*	*Random effects*
CDR	-362.70 (239.51)	-3.71 (65.16)	0.244 (66.66)
NSDPpc	-0.052** (0.023)	0.096** (0.017)	0.075* (0.018)
APPB	-152.91* (10.18)	-4.79* (20.42)	-44.77* (17.80)
PD	12.86 (35.86)	-134.40 (86.98)	-98.48 (70.59)
NF	-0.012 (0.011)	0.017*** (0.019)	0.010 (0.008)
EMP	3.22e-05 (8.38e-05)	-6.36e-06 (2.24e-05)	-2.57e-06 (2.29e-06)
Constant	3437.00* (218.50)	715.11** (302.31)	1237.00* (342.60)
R2-within	-	0.237	0.225
R2-between	-	0.0006	0.195
R2	0.468	0.006	0.230
$\rho = (\sigma_u^2/[\sigma_u^2 - \sigma^2_\varepsilon])$	-	0.966	0.940
Hausman specification test - χ^2 value (p-value):		16.27* (0.002)	
N		290	

Note: Standard deviations in parentheses. *, **, *** significant at 1, 5 and 10 percent levels.

there is random effect or differences between the models are not systematic against the alternative hypothesis that there is no difference is rejected by the significant chi-square test. Hence, the fixed effects model is appropriate for estimating deposit penetration. Thus, per capita NSDP, the proxy for income, and not the population size, level of industrialisation or number of bank branches, is the important determinant of deposit penetration in states of India.

Table 4 presents the empirical estimates of the determinants of credit penetration in states of India. In the pooled OLS regression, NSDP per capita and average population per branch have a significant negative effect on credit penetration in India. Taking into account the heterogeneity among states in panel regressions yields significant positive income and bank branch effects on credit penetration. Unlike deposit penetration, APPB has a positive significant effect on credit penetration at 10 percent level. Further, the impact number of factories on credit penetration is positive and population density negatively affects credit penetration. However, the credit-deposit ratio has no relationship with credit penetration.

The inter-class correlation of the panel model given by the rho value indicates that more than 90 percent of the variance in credit penetration in states is due to differences across panels. Thus, there also exists a wide inequality in access to formal credit for people between states of India. This wide gap in credit penetration is mainly due to poor branch networking and low per capita income. The within and

Table 4: Panel Pooled OLS, Fixed Effects and Random Effects Regression Estimates of Credit Penetration in the States of India

Dependent variable: Credit penetration

Variable	*Pooled OLS*	*Fixed effects*	*Random effects*
CDR	-6.12 (22.66)	5.52 (6.99)	6.49 (7.23)
NSDPpc	-0.005** (0.002)	0.016* (0.002)	0.013* (0.002)
APPB	-13.59* (0.96)	3.80*** (2.19)	-1.62 (1.87)
PD	-0.058 (3.39)	-22.02** (9.33)	-15.38** (7.34)
NF	0.002 (0.001)	0.005* (0.0009)	0.003* (0.0009)
EMP	1.08e-06 (7.93e-06)	-2.34e-06 (2.40e-06)	-1.78e-06 (2.48e-06)
Constant	310.20* (20.68)	-11.17 (32.44)	61.35*** (34.98)
R2-within	-	0.443	0.429
R2-between	-	0.055	0.018
R2	0.468	0.018	0.047
$\rho = (\sigma_u^2/[\sigma_u^2 - \sigma^2_\varepsilon])$	-	0.966	0.926
Hausman specification test χ^2 value (p-value):		23.22* (0.00001)	
N		290	

Note: Standard deviations in parentheses. *, **, *** significant at 1, 5 and 10 percent levels.

between R-square values show that around 42 percent of variations in credit penetration is explained by the differences in NSDP and employment levels within the states of India and only around 5 percent is due to variations in factors across the states in India. The highly significant chi-square value of Hausman's specification test rejects the null hypothesis of random effects and accepts the fixed effects model as the most appropriate model for credit penetration too.

CONCLUSION

Noting that the level of financial inclusion is very low in India and the level of financial activities widely differ across states of India, this study tries to identify the determinants of, and estimate the effects of such factors, on financial inclusion in states of India using state-wise panel data. The indicators of financial inclusion used are deposit and credit penetration, defined as the number of bank accounts of both types per 1000 population. The panel estimation methods of pooled regression, fixed effects regression and random effects regression are used on an unbalanced panel data set for 26 states over 13 years from 2001 to 2013. The analysis shows relatively higher levels of financial inclusion in states like Himachal Pradesh whereas states like Nagaland and Bihar show very low levels of accounts penetration. The panel estimates of deposit and credit penetrations show significant NSDP per capita income effect on deposit penetration, whereas not only income level but also bank branch networking and access to banks determine credit penetration.

The results of this study on the intensity of financial inclusion in states of India also reveal that financial inclusion not only depends on banking sector variables but also other factors that determine financial inclusion in a country. The differences in the intensity of financial inclusion are mainly due to differences in state-level development and economic factors. In case of deposit penetration as a measure of financial inclusion, income and industrialisation of states play a vital role, whereas in the case of credit penetration, apart from income and industrialisation, factors like population density and branch networking matter a lot. There also exists wide variations between states in financial inclusion which requires effective policy initiatives especially on access and availability of formal financial services.

References

Allen, F., A. Demirguc Kunt, L. Klapper and M.S.M. Pería (2016). 'The foundations of financial inclusion: Understanding ownership and use of formal accounts', *Journal of Financial Intermediation*, 27, 1, 1-30.

Alter, A. and B. Yontcheva (2015). 'Financial inclusion and development in the CEMAC', IMF Working Paper No.15/235.

Camara, N., X. Pena and D. Tuesta (2014). 'Factors that matter for financial inclusion: Evidence from Peru", Working paper no. 1409, Economic Research Department, BBVA Bank.

Chakravarty, S.R. and R. Pal (2013). 'Financial inclusion in India: An axiomatic approach', *Journal of Policy Modeling*, 35, 5, 813-837.

Chithra, N. and M. Selvam (2013). 'Determinants of financial inclusion: an empirical study on the inter-state variations in India", SSRN paper no. 2296096.

Dixit, R. and M. Ghosh (2013). 'Financial inclusion for inclusive growth of India - A study of Indian states', *International Journal of Business Management & Research*, 3, 1, 147-156.

Gupta, A., V. Chotia and N.V.M. Rao (2014). 'Financial inclusion and human development: A state-wise analysis from India', *International Journal of Economics, Commerce and Management*, 2, 5, 48-68.

Hausman, J. (1978). 'Specification tests in econometrics', *Econometrica*, 6, 46, 1251-1271.

Inoue, T. and S. Hamori (2012). 'How has financial deepening affected poverty reduction in India: Empirical analysis using state-level panel data', *Applied Financial Economics*, 22, 5, 395-408.

Kumar, N. (2013). 'Financial inclusion and its determinants: Evidence from India', *Journal of Financial Economic Policy*, 5, 1, 4-19.

Marin, A.G. and R. Schwabe (2012). 'Competition and barriers to banking', *Mexican Central Bank Journal*, 95, 14, 20-46.

Nandru, P., B. Anand, and S. Rentala (2016). 'Exploring the factors impacting financial inclusion: Evidence from South India', *Annual Research Journal of Symbiosis Centre for Management Studies*, 4, 1-15.

Reserve Bank of India (various issues) *Handbook of Statistics on Indian Economy*.

Reserve Bank of India (various issues) *Basic Statistical Returns for Commercial Banks*.

Sarma, M. (2008). 'Index of financial inclusion', Working paper no. 215, Indian Council for Research on International Economic Relations.

Sarma, M. (2015). 'Measuring Financial inclusion', *Economics Bulletin*, 35, 1, 604-611.

Sarma, M. and J. Pais (2011). 'Financial inclusion and development', *Journal of International Development*, 23, 5, 613-628.

Tuesta, D., G. Sorensen, A. Haring and N. Camara (2015). 'Financial inclusion and its determinants: The case of Argentina', Working paper no. 1503, Economic Research Department, BBVA Bank.

To cite this article:

T. Lakshmanasamy. Financial Inclusion in the State of India: A Panel Data Analysis of Accounts Penetration. *Journal of Development Economics and Finance*, Vol. 1, No. 2, 2020, pp. 209-225