NET Institute*

www.NETinst.org

Working Paper #07-29

September 2007

Network Economics and the Digital Divide in Rural India

P.D. Kaushik
Rajiv Gandhi Institute of Contemporary Studies

Jake Kendall, Nirvikar Singh, Kristen Williams
UC Santa Cruz

Yan Zhou
California State University, Sacramento

Network Economics and the Digital Divide in Rural India*

P.D. Kaushik↑

Jake Kendall±

Nirvikar Singh±

Kristen Williams±

Yan Zhou↓

September 30, 2007

Abstract

The idea of a 'global digital divide' is well accepted, and cross-country studies of determinants of differences in computer and Internet penetration have identified income, telecommunications infrastructure, and regulatory quality as key influencing factors. The policy implications from these studies are relatively blunt: get richer, have more telephones, and regulate telecommunications better. In this paper, we examine an alternative policy approach to bridging the digital divide, through organizational innovations that provide low cost Internet access in developing countries, within the existing levels of income, telecommunications infrastructure and regulatory environment. We use survey data from 500 individuals in four states of India: Haryana, Madhya Pradesh, Punjab and Rajasthan, to examine factors influencing patterns of computer and Internet use. The situations in which data was collected were ones where computer and Internet access was being provided by a developmental agency (government or non-government). We estimate logit and multinomial logit models, using explanatory variables such as income, household size, education, and occupation, as well as infrastructure factors such as quality of electricity supply, and availability of telephones and televisions. Thus we are able to go beyond simple analyses of penetration at the country level, to understand the microeconomics of computer and Internet use in rural India. In particular, by examining patterns of use, we are able to comment on the importance of network externalities for diffusion of computers and the Internet in these local rural contexts.

JEL Classifications: O1, L86
Keywords: Development, IT, ITC, India, Digital Divide, Internet

---

* We thank the NET Institute (www.NETinst.org) for financial support.
↑ Rajiv Gandhi Institute of Contemporary Studies
± University of California, Santa Cruz
↓ California State University, Sacramento

1 Introduction

The term 'digital divide' refers to inequalities in access and use of digital technologies such as personal computers, the Internet, and even, in the case of developing countries, telephones. Within developed countries, the focus has been on digital divides that line up with inequalities in income, but also factors such as race (e.g. Fairlie, 2004). In contrast to within-country divides, the concept of a 'global digital divide' focuses on inequalities in computer and Internet penetration across countries, particularly looking at differences between developed and developing countries. For example, in a recent, comprehensive study, Chinn and Fairlie (2006) use cross-country regressions to analyze a data set of 161 countries over the period 1999-2001. They find that a major portion of the global digital divide (in terms of levels of computer and Internet use) is explained by income per capita, with telephone density and regulatory quality following in importance as explanatory factors. The Chinn-Fairlie study is distinguished from most previous analyses by providing a ranking of causal factors in terms of importance, as well as by the use of a more recent and broader data set. However, the main conclusions from all such studies are relatively unsurprising, and the policy implications relatively blunt: get richer, have more telephones, and regulate telecommunications better.

An alternative approach to understanding how the digital divide may be reduced, or even bridged, is to examine microeconomic determinants of computer and Internet use, in situations where low cost Internet access in developing countries is provided by government or nongovernmental agencies. In assessing the cost-effectiveness of such efforts, or in designing them to make success more likely, understanding the motivations of individual users, and their patterns of use, is clearly important. In such cases, rather than relying on economic growth or country-level shifts in regulatory policies, the attempt is to overcome the digital divide within the existing conditions of income levels, telecommunications infrastructure and regulatory environment. On the supply side, it can be important to understand the factors that lead to commercial viability of rural Internet kiosks or telecenters (e.g. Kendall and Singh, 2006). On the demand side, it is crucial to understand the characteristics of individuals or households that affect their choices with

respect to using computers and the Internet. This kind of microeconomic analysis provides a more fine grained perspective on the factors that can influence the digital divide, beyond obvious ones such as per capita incomes. Identifying potential barriers to adoption when computer and Internet access are made available may help to generate policy recommendations that are achievable within the existing macro environment and without massive infrastructure investments which may prove unsustainable ex-post.

This paper uses survey data from 500 individuals from locations in four different states of India (Haryana, Madhya Pradesh, Punjab and Rajasthan) to analyze the individual or household-level determinants of computer and Internet use, in situations where access has been provided by a developmental agency (government or non-government).[1] We have data on characteristics such as income, household size, education, and occupation, as well as infrastructure factors such as quality of electricity supply, and availability of telephones and televisions. We also have data on motivations for the use of computers and the Internet, including educational, commercial and governmental purposes. Thus we are able to examine how socio-economic characteristics affect choices of whether to use computers and the Internet, and, if so, for what purposes to use them.

The rest of this paper is organized as follows. In section 2, we provide some further discussion of our work and its contribution to the literature. In particular, we seek to explore microeconomic factors behind the digital divide, as well as implications of rural IT for development and the possible role of network economies of scale. Section 3 describes the data, including variables available, and summary statistics, as well as the econometric techniques used. Our empirical technique relies on logit and multinomial logit to model behavior with respect to computer and Internet use, and broad patterns of usage in each case. Section 4 presents the results of our econometric analysis. Key results are the importance of education and English fluency in supporting computer and Internet use, while income levels, age and gender seem to be less important. There is no evidence of usage patterns that require strong network effects to flourish. Section 5 is a summary conclusion, also outlining suggestions for future research.

[1] Reports on related fieldwork in India include Singh (2004a, 2004b, 2006).

2 Motivation and Literature

This paper begins to fill a key gap in the literature on the global digital divide,[2] by providing a microeconometric analysis of the factors affecting computer and Internet use in rural areas of developing countries, where the majority of the world's population still lives. This research also contributes to the debate on the role of IT in development, which has suffered from the absence of formal econometric studies that provide quantitative estimates of the *microeconomic* impacts and determinants of IT use in the development context.[3] The last two decades have seen the emergence of information technology (IT) as a major force for change in developed countries. While the virtues of IT have sometimes been over-hyped, one can reasonably argue that IT has begun to have a significant impact on the lives of people in developed, industrialized countries. The benefits of IT for developing countries are less clear-cut. IT can improve efficiency, make developing country firms more globally competitive, and bring many benefits to well-off consumers – whose consumption patterns are close to those of the developed world – in these countries. From this perspective, IT is of limited relevance to the poor in developing countries, lacking basic health, sanitation and education. On the other hand, there have been numerous attempts to harness the power of IT in developing countries, to try to improve the delivery of such basic services, as well as provide other services that might have been otherwise inaccessible to poor, isolated villagers. Many case studies, newspaper articles and web sites describe various achievements in the use of IT to improve the lives of the poor in poor countries. While we do not have data to directly measure impacts, we can document the extent and patterns of use, and make some inferences from these patterns.

We can also directly gain insights into the behavior of rural Internet users in developing countries, a group whose choices have not previously been analyzed through formal econometric techniques. These results can have wider implications for

[2] See Chinn and Fairlie (2006) for comprehensive references to the literature.
[3] Microeconomic studies in the context of developed country firms include Brynjolfsson and Hitt (2003) and Bresnahan, Brynjolfsson, and Hitt (2002).

understanding the future spread of e-commerce, and also the overcoming of the global digital divide. By examining the motivations for using computers and the Internet in these cases, we can improve our understanding of demand side barriers to use, which often arise from network economics. For example, if rural Internet users mainly seek information about global market conditions for their crops, individual adoption does not require other local adopters. On the other hand, if the Internet is used for local communication, or commercial transactions, it will be critical that a sufficient number of other local residents are also users. Recent theoretical research (e.g. Ryan and Tucker (2006)) shows that in an environment where agents are heterogeneous and have diverse needs driving demand for ICT services, welfare can be improved by strategically targeting the right subgroup to stimulate adoption. Our work complements this work by measuring the heterogeneity of consumers and their demand needs. We believe our measurements complement network adoption theories to facilitate the formulation of actionable policy directives.

As noted in the introduction, the Chinn and Fairlie (2006) study is the most up-to-date and comprehensive with respect to global digital divide issues.[4] In addition to the variables mentioned that most affect levels of computer and Internet penetration, i.e., income per capita, telephone density and regulatory quality, other factors that matter include per capita electricity consumption up to a threshold level, and human capital variables, which have the right sign but are statistically insignificant. Chinn and Fairlie also find that urbanization rates have a negative impact on Internet penetration, which is explained by Dewan, Ganley and Kraemer (2004) in terms of the availability of substitute means of communication in urban environments, though this may be hard to establish in any definitive manner.

The study by Dewan, Ganley and Kraemer (2004) provides a further dimension of analysis. They use quantile regressions to examine how the impact of individual socio-economic factors varies with the level of IT penetration.[5] Using this technique, they find that not only is the level of IT penetration increasing in GDP per capita, but the influence

[4] Chinn and Fairlie survey and critique many previous cross-country studies on the digital divide, and the reader is referred to their paper for such additional references.

[5] They also examine IT penetration more broadly than simply Internet penetration, including PCs and mainframes, using data over a relatively long period.

of this factor is stronger in countries with higher IT penetration. For mainframes and PCs, telephone lines per capita, years of schooling, and trade in goods are found to have had similar but stronger effects in developed countries than in developing countries. On the other hand, for Internet users, telephone costs, years of schooling and trade in goods have a stronger impact in countries at lower penetration levels, even controlling for wealth differences. This study therefore suggests that the factors influencing the digital divide are quite nuanced, and not fully captured by standard cross-country regressions.

Another approach to cross-country regression results is provided by Mariscal (2005). After performing the regression, she compares actual with fitted values for teledensity, and finds that countries such as Mexico, Guatemala and Peru lag behind what would be predicted by their development levels. In the case of Mexico, she argues that that the digital divide is not narrowing even with the deployment of telecommunications networks: "Even though telecommunications penetration has increased substantially in Mexico during the past 10 years, this country is still well behind other countries with similar development levels and, as has been the case in many countries, the deployment of IT has been very unequally distributed." (Mariscal, 2005, p. 423). This observation ties in with the analysis of our paper, which focuses precisely on IT use by those in developing countries who may be left out even when IT penetration improves at the level of country averages.

3 Data and Empirical Methods

Data

The data was collected in field surveys supervised and conducted by one of the authors (P.D. Kaushik). Residents of each survey region were asked a detailed series of questions pertaining to household socioeconomic characteristics, economic activities engaged in, and patterns of use of ICTs. Each of the four samples was collected in an area where a governmental agency or NGO provided computer and Internet access to the surrounding population via a centrally located internet kiosk. The computer and media usage patterns of respondents are described in Table 1, which presents sample sizes and

some summary statistics. It can be seen from Table 1 that each sample location has a fairly high proportion of computer and Internet users.

Table 1: Data Overview

Initiative	Sample Size	Computer Users (%)	Internet Users (%)	TV Owners (%)	Newspaper Readers (%)	Telephone Owners (%)
JanMitra (Rajasthan)	200	55.5%	55.5%	60%	74.5%	27.5%
Gyandoot (Madhya Pradesh)	200	65%	54.3%	72.4%	77.5%	36%
Drishtee (Haryana)	50	78%	78%	64%	80%	34%
Tarahaat (Punjab)	50	84%	84%	88%	86%	52%

In addition the sample is reasonably well-off, in the sense of having televisions and (less commonly) telephones. Individuals in the sample also rely to a considerable extent on newspapers for information. The picture is one of relatively sophisticated media consumers who might have many different reasons for recourse to the internet.

Table 2 provides summary data on the initiatives that supported Internet access in each location. This consists of a mix of government and nongovernmental organizations.

Table 2: Overview of Sample Digital Divide Initiatives

Initiative	Organizational Basis	Number of Centers	Population Served per Center
JanMitra	Government	20-30	20,000-25,000
Gyandoot	Government	20-30	20,000-25,000
Drishtee	Civil Society – Drishtee Foundation	20-30*	10,000-15,000

Tarahaat	Civil Society – Development Alternatives	8**	20,000+

Notes:. * Sirsa District only – Drishtee has a larger presence in other states
** At time of survey – now there are approximately 40 centers

In addition to information about usage and patterns, we have collected data on various individual or household characteristics. The main variables are summarized in Table 3. Further details of the variables and sample proportions are provided in an Appendix.

Table 3: Main Variables

Dependent variables	Independent variables
Use computer or not	Sex
Purpose of using computer	Marital status
• Educational	Age
• Commercial	Education
• Personal	Main occupation
• Other	Income level
Use Internet or not	Sale or own consumption
Purpose of using Internet	Electricity supply quality
• Educational	Household income
• Commercial	Newspaper readership
• Personal	TV ownership
• Government-related	Telephone ownership

Empirical Methods

The empirical analysis proceeds by using logit and multinomial logit regressions to analyze the choice to use computers, the choice to use the internet, and the choice between various categories of internet activity/service. We analyze each location separately, in order to allow for different local effects and the possibility of different structural relationships in each area. This permits us to make comparisons across locations to isolate similarities and differences in the factors that affect computer and Internet use on the demand side.

We perform logit analysis for each of the binary decisions: whether to use a computer or not, and (of computer users) whether to use the Internet or not. For computer and Internet users, we also estimate multinomial logit specifications, taking into account three possible categories of use (see Table 3) to determine what factors drive individuals' internet needs.

Unfortunately the dummy variables in various categories are not perfectly comparable across survey sites. This was in most cases due to lack of responses in certain categories and multicollinearity concerns (e.g. when no respondents indicated having income greater than a certain level, the high income dummy had to be redefined with a lower cutoff for that site). The variable definitions for all sites and regressions were constructed to be as comparable as possible given these constraints. The exact definitions are given in the appendix.

4 Results

Computer Usage

Table 4 shows results logit regressions for the determinants of computer usage at the four survey sites. While the survey methodology and questions were the same across sites, differences in local conditions in variables such as education and income required us to combine categories differently for the sites in the case of such variables. However, education and income categories are ordered in the same manner across regressions (e.g. the dummy Income 1 indicates less income than Income 2 if they are in the same regression, but may not be comparable levels of income across regressions). The omitted dummy in the cases of education and income is the lowest level in the category.

Though the results are not uniform across sites, some patterns do emerge. Higher levels of education are associated with higher computer use, though this coefficient is not significant in Punjab. Similarly, greater age is associated with less computer use except in Rajasthan.

Interestingly, television ownership has a negative (if insignificant) effect on the use of computers in three of the four sites. In Kendall and Singh (2006) the authors find that games are among the most frequently listed internet uses reported by kiosk owners in

rural India. Many households in our sample may see internet based entertainment as a substitute for TV ownership given a binding budget constraint. Reading an English language newspaper has a positive and marginally significant effect on the decision to use a computer. Similarly, cable connection ownership is also negative and significant in the Madhya Pradesh regression though positive in the Rajasthan regression.

The data lends support to the notion that consumers with high desire for other media and communication services perceive the internet as a complementary good. Telephone usage is positive in the regressions where it is a statistically significant variable (Punjab and Haryana). Most likely, telephones and internet are both used by relatively more sophisticated individuals who have greater communication needs. Further, the coefficient on the variable indicating that the respondent reads English newspapers is also positive and significant in two of three regressions (Madhya Pradesh and Rajasthan).

It appears that computer usage does not differ significantly between men and women. In this sample however 181 men and only 19 women were surveyed so the power of the test is low.

Internet Usage

Table 5 presents results of the internet usage logit regressions. As with the computer usage regressions, the variable definitions are not exactly the same across sites (see Appendix A for exact definitions).

The results of the internet regression tell a slightly stronger story than the computer use regressions. Here being female and less educated both seem to indicate lower recourse for accessing the internet, the parameters on these variables are significant in 2 of three regressions. Again, respondents who read both English and local language newspapers and those who own telephones use the internet more. The negative relationship between TV owners and cable connection owners also holds with the exception of Rajasthan where cable connection ownership is positively and significantly associated with internet use.

Age exhibits a mild positive association with usage of the internet in the regressions where it is a significant variable. The mixed results with this variable across the use computer and use internet regressions may indicate a non-linear relationship – it is likely the youngest and oldest have less recourse to internet and computers. Unfortunately our data do not allow us to test a non-linear specification.

The conclusions drawn from the internet usage regressions are similar to those of the computer usage regressions. We find that media and communication consumption seems to be complementary with internet use with the exception of television ownership, possibly due to a budget constraint. We also find some evidence that women and less educated individuals use the internet less. The profile of internet user in these locations seems to be educated and media sophisticated men. Interestingly, income plays almost not role in internet or computer usage. This is likely due to the fact that internet usage is being subsidized at the sample sites by government or NGO support.

Purposes of Computer Use

These regressions are performed on subsamples of computer users, and the multinomial logit specifications seek to identify how particular characteristics affect the relative likelihood of different choices. Thus, the coefficients for any explanatory variable sum to zero across the four choices, which are educational, commercial, personal and other uses. Personal use includes general browsing, communication, and so on, while 'other' includes government-related activities, such as filling out application forms. In two of the case, Haryana and Punjab, the sample sizes are relatively small, and obtaining convergence required omitting many of the potential explanatory variables, because of insufficient variation within the sample, and multicollinearity problems. Hence, the final specifications presented in Table 6 vary considerably across the four states. The overall fit and explanatory power of the regressions is also quite low. Nevertheless, a few observations are possible.

First, in the case of Haryana, computer users who read an English language newspaper were less likely to do so for educational purposes, in contrast to the case in the Rajasthan location. There are no other significant effects in the case of Haryana, and the

result for newspaper readers for that state may simply be an anomaly. In the case of television ownership, such computer users were more likely to have educational purposes in Punjab, in contrast to Rajasthan, where their use leaned toward commercial motives as their primary reason for computer use. In this case, the difference is understandable when one knows that Internet access was much poorer in the Punjab sample, and that those kiosk operators relied on educational offerings for revenue, while in the Rajasthan case, Internet access was superior.

Owning a telephone skewed computer use away from commercial purposes in both Madhya Pradesh and Rajasthan, possibly reflecting the fact that a telephone is adequate for the most basic commercial purpose, namely price discovery. In fact, it is well-documented that mobile or fixed-line phones are an important source of such information (e.g., Economist, 2005; Jensen, 2005). Other kinds of commercial uses, such as accounting, are not suited to shared-access kiosks, and trading or matching buyers and sellers did not occur in these locations.

We did not find significant gender effects on the pattern of usage, and the only age effect was in the case of the Madhya Pradesh sample, where younger computer users seemed to be less likely to have 'other' uses, which is consistent with these being weighted toward government-related activities. In particular, the Madhya Pradesh site is well known as one of the first attempts in India to provide rural e-governance services. Income, education and computer ownership had no significant effects on the pattern of computer use. A reasonable conjecture, generalizing from the contrast with respect to educational purposes in the cases of Punjab and Rajasthan, is that usage patterns in such cases are more determined by supply conditions, including the quality of the infrastructure, and the nature of the applications made available.[6] Since the foregoing refers to computer usage patterns, the issue of network effects does not arise: for that we turn to patterns of Internet usage.

[6] In a different sample, from kiosks in southern and western India in which a range of applications was available with more homogeneous infrastructure, Kendall and Singh (2006) found that their was a wide variety of uses across individual kiosks, so that demand patterns do differ across locations. However, in that case, data on individual patterns of use were not available.

Purposes of Internet Use

Table 7 presents the results of multinomial logit regressions, which regress the various purposes for which the internet was accessed against the explanatory variables used earlier. Again, there were some problems with convergence of the estimates and in the case of Haryana and Punjab, the small sample sizes meant that some purposes were hardly represented in the sample. Thus, for Haryana we combined educational and personal usage of the Internet, and in Punjab, we combined educational and government-related usage. The Haryana and Punjab data also exhibited the anomaly that Internet usage was reported to be greater than computer usage. Thus, the results for these two states may not be very reliable.

In the case of Internet usage patterns, reading an English newspaper did not have any impact, except in the case of the Haryana sample, where such individuals were more likely to be skewed toward government-related use, and away from commercial purposes, though we offer caution about the robustness of such a conclusion based on that sample. Again, there were no education effects on the pattern of usage in any of the four locations. Nor were age effects on patterns of use discernible: even in the case of Madhya Pradesh the negative impact of age on government-related usage is no longer significant (in contrast to Table 6). The positive connection between television ownership and educational usage in the Punjab sample is present for Internet usage as well, and the same sample reveals a negative impact of income on usage of the Internet for (miscellaneous) personal motives. Finally, the Rajasthan sample indicates that males were more likely to use the Internet for educational purposes, though the relevant coefficient is only significant at the 10 percent level.

Overall, it did not appear that there were strong observable factors influencing the pattern of Internet usage. In particular, there was no evidence that income or education had major impacts on usage patterns, and these patterns were not reflective of network effects in Internet usage. For example, individuals who participated in the market economy, through selling crops or livestock products, and individuals who used the internet for personal purposes (both, activities for which network effects might matter), did not have identifiably different patterns of Internet usage.

5 Conclusions

In this paper, we have used data on 500 individuals across four states in north and central India, to examine factors that influence computer and Internet use. By taking surveys in areas where computer and Internet services have been provided by NGOs and/or governmental agencies, we have been able to examine fairly rich usage data. An advantage of this kind of microeconometric exercise is that one is able to go beyond country-level differences, and focus on more precise behavioral effects. In the analysis undertaken here, access problems have been overcome to some extent, or are fairly homogeneous across individuals in any single location. However, there is still considerable variation in factors such as income and education, and we are also able to examine age and gender effects.

At the local level, income does not seem to be as important a determinant of computer and Internet use as is education. In addition to education, there is also strong evidence that some degree of English language fluency is important for IT usage. This conclusion is consistent with the observation that lack of local language content in developing countries may be an important contributing factor to the digital divide on the demand side, even when supply side constraints are mitigated. We did not find any evidence that age and gender are significant contributors to the digital divide in such local settings, though it is possible that there may be some sample selection biases inherent in a survey conducted by outsiders. Certainly, the samples were weighted toward males in all cases. The policy implications of these results are that education and local language content, combined with low-cost access to ICTs, may go a long way toward overcoming digital divides in developing countries.

When we examined patterns of computer and Internet usage, we found very few significant explanatory factors. In particular, there was no evidence that usage patterns might be skewed or constrained by an inability to benefit from network scale. In essence, rural users in such cases were able to find uses for computers that did not rely on networks at all, and network-based uses appeared to be targeted, or tapping into larger existing networks. Thus, there was no indication of ‘eBay’ type network externalities which might hinder ICT penetration. On the other hand, we cannot rule out the existence

of these types of externalities based on our data. Further research is required to examine still broader samples of rural computer users, with more fine grained observations of usage patterns.

References

Bresnahan, Timothy, Erik Brynjolfsson, and Lorin Hitt, (2002), Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence, *Quarterly Journal of Economics*, February, 339-376.

Brynjolfsson, Erik, and Lorin Hitt (2003), Computing Productivity: Firm-Level Evidence, *Review of Economics and Statistics*, November, 793-808.

Chinn, Menzie D., and Robert W. Fairlie (2006), The determinants of the global digital divide: a cross-country analysis of computer and internet penetration, *Oxford Economic Papers*, 16 – 44

Dewan, Sanjeev, Dale Ganley, and Kenneth L. Kraemer (2004), Across the Digital Divide: A Cross-Country Analysis of the Determinants of IT Penetration, Working Paper, Personal Computing Industry Center, Graduate School of Management, University of California, Irvine, November.

Fairlie, R.W. (2004) Race and the digital divide, *Contributions to Economic Analysis & Policy, The Berkeley Electronic Journals*, 3, Article 15, 1–38.

Economist (2005), The Real Digital Divide, Technology and Development Survey, *The Economist*, March 10th.

Jensen, Robert (2005), The Visible Hand(set): Mobile Phones and Market Performance in South Indian Fisheries – The Micro and Mackerel Economics of Information, presentation to NBER-NCAER conference, Neemrana, Rajasthan, India

Kaushik, P.D. and Nirvikar Singh, (2004) Information Technology and Broad-Based Development: Preliminary Lessons from North India, *World Development*, 32 (4), pp. 591-607

Kendall, Jake and Singh, Nirvikar, (2006) "Internet Kiosks in Rural India: What Influences Success? *NET Institute Working Paper* #06-05

Ryan, Stephen and Tucker, Catherine, (2006) Heterogeneity and Dynamics of Technology Adoption, *NET Institute Working Paper* #06-26.

Singh, Nirvikar, (2004a), Information Technology and Rural Development in India, in *Integrating the Rural Poor into Markets*, in Bibek Debroy and Amir Ullah Khan, eds., New Delhi: Academic Foundation, pp. 221-246

Singh, Nirvikar, (2004b), Transaction Costs, Information Technology and Development, Economics Department working paper, University of California, Santa Cruz, October

Singh, Nirvikar, (2006), Crossing a Chasm: Technologies, Institutions and Policies for Developing a Regional IT Industry, forthcoming in conference volume, ed. V. Baskar and V. Gayathri, Sage Publications

Appendix: Tables

Table 4 **Use Computer - Logit - Marginal Effects**

RHS Variable (std. error)	**Haryana** dy/dx	**Madhya Pradesh** dy/dx	**Punjab** dy/dx	**Rajasthan** dy/dx
Gender	0.340*	-0.205***	-0.084	-0.199*
	{0.179}	{0.062}	{0.190}	{0.107}
Age	-0.012**	-0.003	-0.003	0.007*
	{0.006}	{0.002}	{0.007}	{0.003}
Education 1	0.395***	0.100	0.123	0.598***
	{0.141}	{0.068}	{0.154}	{0.060}
Education 2		0.443***		
		{0.052}		
Income 1		0.040	0.026	0.025
		{0.067}	{0.157}	{0.103}
Electricity Connection				-0.064
				{0.104}
English Newspaper		-0.021	0.329**	0.237**
		{0.141}	{0.133}	{0.109}
Own Television	-0.146	0.181	-0.071	-0.151
	{0.173}	{0.137}	{0.205}	{0.121}
Own Cable Connex.		-0.172*		0.203*
		{0.099}		{0.112}
Own Telephone	0.389***	0.024	0.259*	-0.213*
	{0.151}	{0.075}	{0.151}	{0.112}
Observations	50	199	50	200
Pseudo-R-squared	0.3081	0.2097	0.1342	0.2964

*Notes: Standard errors in parenthesis; *** p<0.01, ** p<0.05, * p<0.1*
Due to insufficient sample variation, education, income, and occupation dummies are defined slightly differently across provinces (see Appendix A). Lowest category is omitted when applicable (i.e. no education, low income)

Table 5 **Use Internet - Logit - Marginal Effects**

RHS Variable (std. error)	**Haryana** dy/dx	**Madhya Pradesh** dy/dx	**Punjab** dy/dx	**Rajasthan** dy/dx
Gender	0.151	-0.317***	-0.101	-0.195*
	{0.146}	{0.074}	{0.079}	{0.106}
Age	-0.002	-0.004	0.009*	0.007**
	{0.004}	{0.003}	{0.004}	{0.003}
Education 1		0.173**		0.606***
		{0.087}		{0.061}
Education 2		0.615***		
		{0.051}		
Income 1	-0.063		-0.107	
	{0.125}		{0.091}	
Occupation 1	-0.119	-0.073		0.133
	{0.123}	(0.088)		(0.084)
English Newspaper				0.247**
				(0.103)
Own Television	-0.188*	0.182		-0.117
	{0.110}	(0.149)		(0.108)
Own Cable Connex.		-0.136	0.076	0.198*
		(0.123)	(0.191)	(0.107)
Own Telephone	0.120	0.024		-0.137
	{0.120}	(0.094)		(0.107)
Observations	50	198	50	200
Pseudo-R-squared	0.1379	0.3080	0.1084	0.3072

*Notes: Standard errors in parenthesis; *** p<0.01, ** p<0.05, * p<0.1*

Due to insufficient sample variation, education, income, and occupation dummies are defined slightly differently across provinces (see Appendix A). Lowest category is omitted when applicable (i.e. no education, low income)

Table 6 Purposes of Computer Use - Logit - Marginal Effects (dy/dx)

RHS Variable	Haryana				Madhya Pradesh				Punjab				Rajasthan			
{std. error}	Educ.	Comm.	Personal	Other	Educ.	Comm.	Personal	Other	Educ.	Comm.	Personal	Other	Educ.	Comm.	Personal	Other
Gender					0.024	0.066	-0.059	-0.031								
					{0.075}	{0.098}	{0.080}	{0.080}								
Age					0.003	0.003	0.000	-0.006**	0.003	0.002	-0.007	0.002				
					{0.002}	{0.003}	{0.003}	{0.003}	{0.003}	{0.002}	{0.011}	{0.011}				
Education 1					0.049	0.051	-0.058	-0.041	0.061	0.010	-0.061	-0.009	0.034	-0.088	0.015	0.039
					{0.076}	{0.096}	{0.072}	{0.073}	{0.058}	{0.030}	{0.250}	{0.250}	{0.106}	{0.100}	{0.083}	{0.059}
Income 1									-0.010	0.034	0.035	-0.057	-0.009	-0.012	-0.038	0.060
									{0.051}	{0.035}	{0.270}	{0.272}	{0.132}	{0.123}	{0.097}	{0.091}
English Newspaper	-0.166*	-0.083	0.251	-0.002									0.124	-0.226*	-0.156**	0.257
	{0.095}	{0.256}	{0.331}	{0.260}									{0.185}	{0.106}	{0.073}	{0.192}
Own Television	-0.001	-0.098	0.004	0.094					0.167*	0.110	-0.287	-0.010	-0.096	0.240*	0.120	-0.264
	{3,053}	{1,554}	{976}	{522}					{0.094}	{0.072}	{0.348}	{0.351}	{0.192}	{0.131}	{0.111}	{0.200}
Own Cable Connex.													0.163	-0.252**	-0.071	0.161
													{0.153}	{0.119}	{0.100}	{0.135}
Own Telephone	0.000	0.401	-0.235	-0.166	0.049	-0.153*	0.020	0.084					0.146	-0.195**	0.065	-0.015
	{503.6}	{697.1}	{0.256}	{0.218}	{0.074}	{0.090}	{0.072}	{0.076}					{0.124}	{0.099}	{0.101}	{0.072}
Own Computer									-0.017	-0.031	-0.121	0.170				
									{0.049}	{0.030}	{0.296}	{0.300}				
Obs.	30	30	30	30	130	130	130	130	31	31	31	31	111	111	111	111
Pseudo-R-squared	0.09	0.09	0.09	0.09	0.03	0.03	0.03	0.03	0.13	0.13	0.13	0.13	0.06	0.06	0.06	0.06

*Notes: Standard errors in parenthesis; *** p<0.01, ** p<0.05, * p<0.1*

Due to insufficient sample variation, education, income, and occupation dummies are defined slightly differently across provinces (see Appendix A). Lowest category is omitted when applicable (i.e. no education, low income)

Table 7 Purposes of Internet Use - Logit - Marginal Effects (dy/dx)

RHS Variable	Haryana			Madhya Pradesh				Punjab			Rajasthan			
{std. error}	Educ. & Personal	Commercial	Govt.	Education	Comm.	Personal	Govt.	Educ. & Govt.	Comm.	Personal	Education	Comm.	Personal	Govt.
Gender	-0.552	-0.001	0.552	0.020	0.115	-0.056	-0.080	0.183	0.030	-0.213	0.100*	0.024	-0.077	-0.047
	{26072}	{12760}	{13311}	{0.037}	{0.088}	{0.082}	{0.109}	{0.250}	{0.195}	{0.229}	{0.061}	{0.098}	{0.110}	{0.129}
Age	-0.004	0.000	0.004	0.001	0.003	0.000	-0.004	-0.002	-0.005	0.007	0.003	-0.001	0.001	-0.003
	{404.6}	{958.66}	{554.03}	{0.001}	{0.003}	{0.003}	{0.003}	{0.008}	{0.007}	{0.005}	{0.002}	{0.003}	{0.003}	{0.003}
Education 1	-0.159	-0.001	0.160					0.072	0.000	-0.072				
	{11000}	{17134}	{6134}					{0.161}	{0.143}	{0.110}				
Income 1	0.665	-0.002	-0.664					0.129	0.189	-0.318**				
	{8593}	{27108}	{35701}					{0.170}	{0.124}	{0.139}				
English Newspaper	-0.084	-0.345***	0.429**											
	{0.167}	{0.124}	{0.197}											
Own Television	0.534	0.001	`-0.535**					0.481**	-0.420	-0.061	0.077	0.018	-0.022	-0.072
	{22615}	{20079}	{0.243}					{0.250}	{0.289}	{0.177}	{0.063}	{0.087}	{0.088}	{0.110}
Own Telephone	0.297	-0.003	-0.294	-0.040	-0.031	-0.086	0.157				0.006	-0.056	0.023	0.028
	{6929}	{59882}	{52953}	{0.037}	{0.087}	{0.071}	{0.100}				{0.071}	{0.180}	{0.114}	{0.141}
Observations	39	39	39	99	99	99	99	42	42	42	91	91	91	91
Pseudo-R-squared	0.3418	0.3418	0.3418	0.0389	0.0389	0.0389	0.0389	0.2021	0.2021	0.2021	0.0308	0.0308	0.0308	0.0308

*Notes: Standard errors in parenthesis; *** p<0.01, ** p<0.05, * p<0.1*

Due to insufficient sample variation, education, income, and occupation dummies are defined slightly differently across provinces (see Appendix A). Lowest category is omitted when applicable (i.e. no education, low income)

Appendix A

Variable summary statistics and definitions by survey site.

Table A.1 Haryana

Variable Name	Variable	Freq./Ave	Codes	Description
	LHS Variables			
Computer Use	use_comp	60%	Yes : 1	Did they use a computer?
		40%	No: 0	
Purpose of Computer	purp_use	16.67%	Education/student : 1	What is the purpose of using the computer?
		36.67%	Commercial : 2	
		26.67%	Personal / official : 3	
		20.00%	Others : 4	
Internet Use	use_inet	78%	Yes : 1	Do they use internet?
		22%	No : 0	
Purpose of Internet	purp_inet	2.56%	Education/student : 1	What is their purpose for using internet?
		28.21%	Commercial : 2	
		25.64%	Personal / official : 3	
		43.59%	Government : 4	
	purp_inet_1	28.21	Educ/Pers/Other : 1	Purposes 1 and 3 were combined here for Haryana
		28.21	Commercial : 2	
		43.59	Government : 4	

RHS Variables

Age	age	Ave age: 37.26 Range: 14-75 yrs		Age of the respondent & other members of the family
	educ	6%	Illiterate : 0	Education level of the respondent & other members of the family
		22%	Semi-Literate : 1	
		14%	Primary : 2	
		16%	Middle school : 3	
		16%	Secondary school : 4	
		16%	Higher secondary : 5	
		6%	Graduate : 6	
		4%	Post graduate : 7	
			Others : 8	
Education 1	educ3_4	42%		combined secondary, higher secondary, graduate and post graduate
Electricity Connection	en_conn	100%	Yes :1 No :0	Does the user have electricity connection?
Income 1	in_come	32%	< 10,000 :1	Household income per annum
		38%	10,000-25,000 :2	
		24%	25,000-50,000 :3	
		6%	> 50,000 :4	
	inc_234	42%		users in the income levels 2,3,&4 combined
	npap_lang	68%	Hindi/Local : 1	Language of the newspaper
English Newspaper		0%	English : 2	

		12%	Both : 1,2	
		0%	None : 0	
	npap_both	12%		
	eng_both_npap_lang	12%		combined households that receive english and both languages
	occ_main		None : 0	Main occupation of the respondent
		42%	Farming : 1	
		6%	Animal Rearing : 2	
		20%	Trade/Business : 3	
		20%	Employed : 4	
		6%	Student : 5	
		6%	Other : 6	
Occupation 1	occ2_farm_anrear	48%		combined occupations 1&2
Own Cable Connection	own_cc	60%	Yes : 1	Do they own cable connection?
		40%	No : 0	
Own Computer	own_comp	8%	Yes : 1	Do they own a computer?
		92%	No : 0	
Own Telephone	own_tel	34%	Yes : 1	Do they own a telephone?
		66%	No : 1	
Own Television	own_tv	64%	Yes : 1	Do they own TV?
		36%	No : 0	
Gender	sex	68%	Male : 1	Sex of the respondent & other members of the family
		32%	Female : 0	

Table A.2 Madhya Pradesh

Variable Name	Variable	Freq/Ave	Codes	Description
	LHS Variables			
Computer Use	use_comp	65%	Yes : 1	Did they use a computer?
		35%	No: 0	
Purpose of Computer	purp_use	18.46%	Education/student : 1	What is the purpose of using the computer?
		43.08%	Commercial : 2	
		17.69%	Personal / official : 3	
		20.77%	Others : 4	
Internet Use	use_inet	`54.27%	Yes : 1	Do they use internet?
		`45.73%	No : 0	
Purpose of Internet	purp_inet	10.19%	Education/student : 1	What is their purpose for using internet?
		25.00%	Commercial : 2	
		13.89%	Personal / official : 3	
		50.93%	Government : 4	
	RHS Variables			
Age	age	Aveage: 43.65%		Age of the respondent & other members of the family
		Range: 16-85		
	educ	`9.5%	Illiterate : 0	Education level of the respondent & other members of the family
		`21.5%	Semi-Literate : 1	
		20%	Primary : 2	
		20%	Middle school : 3	
		8%	Secondary school : 4	
		`12.5%	Higher secondary : 5	

		7%	Graduate : 6	
		`0.5%	Post graduate : 7	
		1%	Others : 8	
Education 1	educ3_second_highsec	`20.5%		combined primary and middle education
Education 1	educ4_grad_postgrad	`7.5%		
	educ3_4	28%		combined secondary, higher secondary, graduate and post graduate
Electricity Connection	en_conn	98%	Yes :1	Does the user have electricity connection?
		2%	No :0	
		5%	Good :1	Combined electricity status 1,2,&3 vs. 4
		46.50%	Avg :2	
		46.50%	Poor :3	
		2%	Unconnected :4	
	in_come	60%	< 10,000 :1	Household income per annum
		33%	10,000-25,000 :2	
		`6.5%	25,000-50,000 :3	
		`0.5%	> 50,000 :4	
Income 1	inc_234	40%	Yes: 1	users in the income levels 2,3,&4 combined
	npap_lang	68%	Hindi/Local : 1	Language of the newspaper
		`2.5%	English : 2	
		7%	Both :1,2	
		`22.5%	None:	
	eng_both_npap_lang	`9.5%		household receives both hindi/local and english language
English Newspaper	npap_both	7%		combined households that receive any english n.p.
	occ_main	0	None : 0	Main occupation of the respondent
		40%	Farming : 1	
		8%	Animal Rearing : 2	
		22%	Trade/Business : 3	

		`15.5%	Employed : 4	
		3%	Student : 5	
		`11.5%	Other : 6	
Occupation 1	occ2_farm_anrear	48%		combined occupations 1&2
		52%		
Own Cable Connection	own_cc	`60.3%	Yes : 1	Do they own cable connection?
		`39.70%	No : 0	
Own Computer	own_comp	6%	Yes : 1	Do they own a computer?
		94%	No : 0	
Own Telephone	own_tel	36%	Yes : 1	Do they own a telephone?
		64%	No : 1	
Own Television	own_tv	`72.36%	Yes : 1	Do they own TV?
		`27.64%	No : 0	
Gender	sex	`77.5%	Male : 1	Sex of the respondent & other members of the family
		`22.5%	Female : 0	

Table A.3 Punjab

Variable Name	Variable	Freq/Ave	Codes	Description
	Left Hand Side Variables			
Computer Use	use_comp	62%	Yes : 1	Did they use a computer?
		38%	No: 0	
Purpose of Computer	purp_use	19.35%	Education/student : 1	What is the purpose of using the computer?
		12.90%	Commercial : 2	
		32.26%	Personal / official : 3	
		35.48%	Others : 4	
Internet Use	use_inet	84%	Yes : 1	Do they use internet?
		16%	No : 0	
Purpose of Internet	purp_inet	40.48%	Education/student : 1	What is their purpose for using internet?
		19.05%	Commercial : 2	
		19.05%	Personal / official : 3	
		21.43%	Government : 4	
	purp_inet 1&4 combined	61.9		
	Right Hand Side Variables			
Age	age	Ave: 47.7 yrs Range: 21-77		Age of the respondent & other members of the family
	educ	10%	Illiterate : 0	Education level of the respondent & other members of the family
		20%	Semi-Literate : 1	
		8%	Primary : 2	
		16%	Middle school : 3	
		8%	Secondary school : 4	

		10%	Higher secondary : 5	
		20%	Graduate : 6	
		8%	Post graduate : 7	
			Others : 8	
Education 1	educ2_prim_midd	24%		combined primary and middle education
Education 1and 2	educ3_4	46%		combined secondary, higher secondary, graduate and post graduate
Electricity Connection	en_conn	100%	Yes :1	Does the user have electricity connection?
		0	No :0	
		40%	Good :1	Combined electricity status 1,2,&3 vs. 4
		40%	Avg :2	
		20%	Poor :3	
		0	Unconnected :4	
	in_come	42%	< 10,000 :1	Household income per annum
		34%	10,000-25,000 :2	
		12%	25,000-50,000 :3	
		12%	> 50,000 :4	
Income 1	inc_234	58%		users in the income levels 2,3,&4 combined
	npap_lang	58%	Hindi/Local : 1	Language of the newspaper
		0%	English : 2	
		26%	Both: 1,2	
		16%	None: 0	
	npap_both	26%		household receives both hindi/local and english language
English Newspaper	eng_both_npap_lang	26%		combined households that receive english and both languages
	occ_main		None : 0	Main occupation of the respondent
		60%	Farming : 1	
			Animal Rearing : 2	
		20%	Trade/Business : 3	
		18%	Employed : 4	
			Student : 5	

		2%	Other : 6	
Occupation 1	occ2_farm_anrear	60%		combined occupations 1&2
Own Cable Connection	own_cc	88%	Yes : 1	Do they own cable connection?
		12%	No : 0	
Own Computer	own_comp	18%	Yes : 1	Do they own a computer?
		82%	No : 0	
Own Telephone	own_tel	52%	Yes : 1	Do they own a telephone?
		48%	No : 1	
Own Television	own_tv	88%	Yes : 1	Do they own TV?
		12%	No : 0	
Gender	sex	84%	Male : 1	Sex of the respondent & other members of the family
		16%	Female : 0	

Table A.4 Rajasthan

Variable Name	Variable	Freq/Ave	Codes	Description
	LHS Variables			
Computer Use	use_comp	55.50%	Yes : 1	Did they use a computer?
		44.50%	No: 0	
Purpose of Computer	purp_use	38.74%	Education/student : 1	What is the purpose of using the computer?
		31.53%	Commercial : 2	
		18.02%	Personal / official : 3	
		11.71%	Others : 4	
Internet Use	use_inet	55.50%	Yes : 1	Do they use internet?
		44.50%	No : 0	
Purpose of Internet	purp_inet	13.19%	Education/student : 1	What is their purpose for using internet?
		19.78%	Commercial : 2	
		18.68%	Personal / official : 3	
		48.35%	Government : 4	
	RHS Variables			
Age	age	Ave: 42.54 Range: 12-84		Age of the respondent & other members of the family
	educ	10.50%	Illiterate : 0	Education level of the respondent & other members of the family
		12%	Semi-Literate : 1	
		19%	Primary : 2	
		17%	Middle school : 3	
		14.50%	Secondary school : 4	
		16%	Higher secondary : 5	

		6.50%	Graduate : 6	
		4%	Post graduate : 7	
		0.50%	Others : 8	
Education 1	educ2_prim_midd	36%		combined primary and middle education
Education 2	educ3_4	41%		combined secondary, higher secondary, graduate and post graduate
Electricity Connection	en_conn	80.50%	Yes :1	Does the user have electricity connection?
		19.50%	No :0	
		14%	Good :1	Combined electricity status 1,2,&3 vs. 4
		27%	Avg :2	
		39.50%	Poor :3	
		19.50%	Unconnected :4	
	in_come	54%	< 10,000 :1	Household income per annum
		36%	10,000-25,000 :2	
		8.50%	25,000-50,000 :3	
		1.50%	> 50,000 :4	
Income 1	inc_234	46%		users in the income levels 2,3,&4 combined
	npap_lang	`65.5%	Hindi/Local : 1	Language of the newspaper
		0	English : 2	
		9%	Both: 1,2	
		`25.5%	None : 0	
	npap_both	9%		household receives both hindi/local and english language
English Newspaper	eng_both_npap_lang	9%		combined households that receive english and both languages
	occ_main	3%	None : 0	Main occupation of the respondent
		41.50%	Farming : 1	
		1%	Animal Rearing : 2	
		13%	Trade/Business : 3	
		32.50%	Employed : 4	
		0	Student : 5	
		9%	Other : 6	

Occupation 1	occ2_farm_anrear	42.50%		combined occupations 1&2
Own Cable Connection	own_cc	36%	Yes : 1	Do they own cable connection?
		64%	No : 0	
Own Computer	own_comp	9.50%	Yes : 1	Do they own a computer?
		90.50%	No : 0	
Own Telephone	own_tel	27.50%	Yes : 1	Do they own a telephone?
		72.50%	No : 1	
Own Television	own_tv	60%	Yes : 1	Do they own TV?
		40%	No : 0	
Gender	sex	84.50%	Male : 1	Sex of the respondent & other members of the family
		15.50%	Female : 0	