NBER WORKING PAPER SERIES

HUMAN CAPITAL AND THE INDIAN SOFTWARE INDUSTRY

Ashish Arora
Surendra K. Bagde

Working Paper 16167
http://www.nber.org/papers/w16167

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2010

This paper draws from the Bagde's PhD dissertation at Carnegie Mellon University. This research was supported by the Sloan Software Industry Center at Carnegie Mellon University. We are grateful to Dennis Epple, Lowell Taylor and Bill Vogt for very helpful comments. The customary disclaimers apply. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Human capital and the Indian software industry
Ashish Arora and Surendra K. Bagde
NBER Working Paper No. 16167
July 2010
JEL No. I2,I28,J24,O15,O53

ABSTRACT

Though previous studies have noted the role of skilled labor in the growth of the Indian software industry, they have not empirically investigated its importance. In this study we study the effect of the supply of engineers, measured by engineering baccalaureate capacity, on the regional growth of the software exports between 1990 and 2003. We find significant effect of engineering baccalaureate capacity on the growth of software exports even after controlling for other relevant factors. This conclusion is especially interesting because much of this capacity is due to private, rather than publicly supported colleges, and testifies to the private willingness to invest in human capital even in poor countries.

Ashish Arora
Fuqua School of Business
Duke University
Box 90120
Durham, NC 27708-0120
and NBER
ashish.arora@duke.edu

Surendra K. Bagde
Heinz College
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh PA 15213-3890
sbagde@andrew.cmu.edu

Non Technical Summary

The importance of human capital -- skilled and creative workers -- to a "high-tech" industry is routinely acknowledged but often public policy discussions tend to focus on more trendy prescriptions such technology parks, venture capital, incubators and university industry centers. Software relies intensively upon human capital, perhaps more so than any other high-tech industry. Software services, the engine of the Indian software sector, is arguably even more human capital intensive than software products. Thus, few would question the role of human capital stocks in the rise of the Indian software industry. Indeed, one might argue that the plenitude of engineers has created a comparative advantage for India in software service exports.

Between 1985 and 2003, undergraduate engineering baccalaureate capacity increased from about 45,000 (59 per million) to about 440,000 (405 per million), even as the total population increased from 765 million to 1086 million. What is less clearly appreciated is that there are significant variations across Indian states in stocks of the relevant human capital, engineers, and that these differences have played an important part in conditioning where the software industry has flourished. Even less well understood are the reasons for this regional disparity in human capital stocks.

In this paper we empirically investigate how software exports by the fourteen major states of India are conditioned by local levels of human capital, as measured by the state level engineering baccalaureate capacity. Our research covers fourteen states of India, for the period 1990-2003[2]. These fourteen states accounted for 83.47 percent of the country's population in 2003, 78 percent geographic area of India, and 79.2 percent of the net domestic product in 2001-02.[3]

We find that differences in software exports by states are related to the supply of human capital even after controlling for factors such as how rich or large the state is, and measures of industrial production, electronics production or telecommunication investment. Since engineering education has been controlled and, in the main, provided by state funded colleges, differences in the willingness of states to invest in engineering colleges could, but do not, explain the bulk of the inter-state variation. Instead, it is the role of private engineering colleges which is the key the puzzle. Simply put, states which allowed private engineering colleges to

[2] Uttar Pradesh and Madhya Pradesh reflect the geographic boundaries as in 1990 and not the current boundaries.
[3] The source for population and area share is the Census of India, and for the GDP share is Government of India, Ministry of Statistics and Programme Implementation (http://mospi.nic.in/9_nsdp_const_9394ser.htm)

enter early were able to get a head start and, this early advantage has persisted for nearly a decade and a half.

In other words, states were favored locations for software development because they had higher stocks of human capital, and they had higher stocks of human capital because they allowed private engineering colleges to operate earlier than other states. We test this hypothesis using a new, hand collected data set of state level software exports and state level engineering baccalaureate capacity in India for a fourteen year period that coincides with the rise of India as a software power. Our identifying assumption is that initially demand for engineers from the software industry was small, and changes in the number of engineers produced was independent of the current or anticipated growth of the software industry. Indeed, a key source of variation is that some states in India allowed the entry of private engineering colleges much earlier than the rest. These tend to be the states that also subsequently became the major poles of software exports.

The number of engineering colleges in India increased from 246 in 1987 to 353 in 1995 and over 1100 in 2003. Eighty percent of new colleges added between 1987 and 1995 were in the private sector and the share of private colleges was even higher at 94 percent for colleges added between 1995 and 2002. Permitting privately financed colleges helped mitigate the adverse effects of the lack of public investments in higher education. It did not completely ameliorate the problem because, as noted earlier, there has been a marked fall in the production of engineering PhDs, even as baccalaureate capacity has increased. Knowledgeable observers of the Indian software industry, and the leading firms themselves, are increasingly concerned about the divergence, which also points to the limits of relying solely upon private financing for human capital development.

Introduction

The Indian software industry, which was almost non-existent till late 1980s, grew at tremendous pace after early 1990s. The Indian software exports were about $128 millions in 1990-91 and grew to $485 millions by 1994-95. By 2003-04 the software exports had increased to $12.2 billions. Though differences in definition imply that US government figures show much lower level of software exports from India, there is no denying that they have grown dramatically over the last two decades, growing at an average of about 30% per year.

India is not the only country to have succeeded in software exports. Israel and Ireland are two other countries that have also achieved software success (Arora and Gambardella, 2005). One common element in all three countries is the role of human capital supply. The importance of skilled manpower, of engineers in particular, to Indian software exports is widely recognized (e.g., Lakha, 1994; Arora and Athreye, 2002). Between 1985 and 2003, undergraduate engineering baccalaureate capacity increased from about 45,000 (59 per million) to about 440,000 (405 per million), even as the total population increased from 765 million to 1086 million. Indeed, one might argue that the plenitude of engineers has created a comparative advantage for India in software service exports.

But the unasked question is – whence did plenitude arise? How did a poor country, with a perpetual problem of financing public expenditure, create so many engineers? An important part of the answer is that it was the private sector that fuelled this increase. For institutional reasons, most of the additional engineering education capacity created in India was in the form of new colleges, and the vast bulk of these colleges were private colleges, privately financed principally from student tuition revenues. The number of engineering colleges in India increased from 246 in 1987 to 353 in 1995 and over 1100 in 2003. Eighty percent of new colleges added between 1987 and 1995 were in the private sector and the share of private colleges was even higher at 94 percent for colleges added between 1995 and 2002. Software exports were also growing rapidly during this period, and software exports and engineering education capacity appear closely linked.

One way to study this relationship is through its regional dimension. Arora et al. (2001) suggest that large share of south and west region in engineering baccalaureate capacity, spurred by the growth of private engineering colleges, was one of main reasons for growth of software industry in those regions. In this paper we empirically investigate how software exports by the fourteen major states of India are conditioned by local levels of human capital, as measured by the state level engineering baccalaureate capacity. The simple point of the paper is that some states were favored locations for software because they had higher stocks of human capital, and they had higher stocks of human capital because they allowed private engineering colleges to operate earlier than other states. We test this hypothesis using a new, hand collected data set of state level software exports and state level engineering baccalaureate capacity in India for a fourteen year period that coincides with the rise of India as a software power. Our identifying assumption is that initially demand for engineers from the software industry was small, and changes in the number of engineers produced was independent of the current or anticipated growth of the software industry.

Our research covers fourteen states of India, for the period 1990-2003[4]. These fourteen states accounted for 83.47 percent of the country's population in 2003, 78 percent geographic area of India, and 79.2 percent of the net domestic product in 2001-02.[5]. As well, the available data require that our measure of software exports is a broad one, including not only the export of software, but also affiliated IT services, including the so-called IT enabled service. IT enabled services are relatively unimportant in terms of revenues for much of our sample and only become significant after the turn of the century. As well, such services, particularly low-end services such as call centers, rarely employ engineers. Their exclusion would therefore only strengthen our results. These, and other issues related to our data are discussed in more detail in the appendix.

There are empirical challenges in such research. First, there may be unobservable state characteristics that are correlated with software industry and engineering baccalaureate capacity. Second, there is issue of endogeneity of engineering baccalaureate capacity. We address this by using a panel dataset and controlling for state fixed effects. We also develop an instrument, discussed in greater detail below, for engineering baccalaureate capacity.

The remainder of this paper is organized as follows. Section 2 briefly reviews the relevant literature dealing with the Indian software industry as well as the literature on agglomeration. Section 3 describes the character, size and regional spread of software exports industry. Section 4 describes the size, growth of technical education in India, its regional dimension and importance of private sector. Section 5 develops a simple model linking the baccalaureate capacity and software production and motivates the empirical specification. Section 6 describes the data. We present results and alternative explanations in section 7. Section 8 summarizes the policy implications and concludes.

---

[4] Uttar Pradesh and Madhya Pradesh reflect the geographic boundaries as in 1990 and not the current boundaries.
[5] The source for population and area share is the Census of India, and for the GDP share is Government of India, Ministry of Statistics and Programme Implementation (http://mospi.nic.in/9_nsdp_const_9394ser.htm)

II Literature review

The first contribution of our paper is to document and explore the role of education policy in the growth of Indian software exports. Specifically, it points to the importance of a human capital producing sector that responds to market demands. In India, access to engineering education was rationed for several years because expanding engineering education capacity was a lower priority for state and federal governments. The slow economic growth in the 1970s and 1980s meant that the social return to such investments was thought to be low. The private return, however, was high, especially for those engineers that went to work overseas. Indeed, it is widely believed that there was excess demand for admission to engineering colleges, with periodic political and legal battles about the ability of privately run colleges to charge market level tuitions.[6] A key policy innovation was to allow privately funded colleges to satisfy this latent demand. As the success of this policy became clearer, other states followed suit. This variation in timing is the key variation that identifies the effect we seek to estimate.

Our findings also speak to the issue of high tech clusters. The bulk of Indian software industry is concentrated in a few clusters; indeed Bangalore has often been branded as the Silicon Valley of India in press accounts. Following Marshall, Ellison, Glaeser and Kerr (2007) argue that ultimately, firms agglomerate to save the costs of transporting either goods (inputs and outputs), people, or ideas, and find support for all three.[7] The literature on clusters and agglomeration is huge and we refer the reader to overviews such as Fujita and Thisse (1996), Fujita, Krugman and Venebles (2001), and Rosenthal and Strange (2004). Our paper is not about clustering or agglomeration in general, but rather about the agglomeration of the Indian software industry. Thus, we are interested in understanding not simply whether and why Indian software exporting firms cluster, but principally *where* they cluster and the reasons for that.

In seeking to understand where high tech industries cluster, stories inspired by the Silicon Valley experience stress the role of an anchor university, labor mobility, venture capitalists and networks of specialized firms (cf. Saxenian, 1994). Others highlight the superior availability of infrastructure (Kapur, 2002). In 1990s many state governments in India set up information technology (IT) parks, which provided physical infrastructure such as building space and electrical power.[8] Though infrastructure was undoubtedly important, our results suggest

---

[6] Since tuitions were regulated and fixed by the government, private colleges circumvented this by charging "capitation fees", which were, in effect, the capitalized value of the tuition they would have liked to charge.

[7] Ellison, Glaeser and Kerr (2007) use patterns of co-agglomeration to quantify the importance of these three transaction costs. Audretch and Feldman (1996) seek to explain the extent to which R&D intensive industries are more likely to agglomerate.

[8] The Software Technology Parks (STP) scheme in 1991 provided reliable internet connectivity and single window clearance for various government permissions to software export firms. There were other schemes like export

that software exports are regionally agglomerated in significant measure because of clustering in a key input, namely engineers.

Klepper (2007) argues that industries cluster in regions which are home to the early industry leaders. He argues that the semiconductor producers cluster in the Silicon Valley because many are spin-offs of some of the early leaders. In particular, he documents that Fairchild, a firm that pioneered the planar process and the integrated circuit, two of the fundamental semiconductor innovations, spawned a number of the firms that would later dominate the Silicon Valley and the semiconductor industry, including Intel, AMD, National Semiconductors, Micron Technology and VLSI Technology. Interestingly, Klepper (2007) argues that a similar process explains the concentration of the automobile industry around Detroit and the tire industry around Akron Ohio. Our results are consistent with Klepper's explanation for industrial agglomeration, but our focus is not on the identities of the firms.

Berry and Glaeser (2006) develop a theoretical model to explain their finding that US metropolitan areas have diverged in terms of skill intensity over time. In their model, entrepreneurs arise randomly from among those with high human capital. These entrepreneurs are assumed to create firms in the cities where they live. If these firms are disproportionately likely to hire high human capital workers, then if workers are mobile, over time cities with higher initial levels of human capital will also disproportionately attract high human capital workers, leading to a divergence across cities in the share of high human capital workers. We too develop a simple model that motivates our empirical analysis. As in Berry and Glaeser, graduating engineers in our model have varying levels of preference for the city in which they go to college (which, in India, is often close to their birthplace). We ignore amenities and assume that firms are price takers in both the product and the input market. It follows that cities with more graduating engineers will attract or create more firms, resulting in these cities having higher software exports.

There are alternative explanations offered for the rise of the software industry in India, some of which have implications for its regional location. The first is the presence of public sector R&D units. Balakrishnan (2006), for instance, has argued that the presence of public sector R&D institutions, including nine defense related laboratories, in Bangalore accounts for the location of the software industry in Bangalore. Once in place, increasing returns would reinforce the initial lead. In our estimations, state fixed effects should account for such

---

processing zones which offered similar incentives to firms locating in such zones. However, STP scheme offers much higher level of flexibility to firms in their location choices and was targeted to software export firms. Firms could locate anywhere and were required to register with designated STP office to avail various incentives.

differences. We also control for electronics production to take into account potential knowledge spillovers. Moreover, as we show below, the factual premise, namely that entrants were initially mostly located in Bangalore is simply wrong.

Srinivasan (2006) singles out telecommunication reforms and the creation of Software Technology Parks (STPs) as key pieces of the puzzle. In our empirical analysis, we control for industrial production and teledensity to control for physical and telecommunication infrastructure. Further, initial software exports consisted of software programmers being sent overseas on short term assignment, and physical and telecommunication infrastructure was less critical than when such work began to be shifted to India. We cannot control for STPs directly. However, as we discuss below, STPs were not important locations for software export prior to 1998 or so.

Other explanations have to do with the role of the diaspora, and the role of entrepreneurship. As we argue in greater detail later, these are complementary to the human capital based explanation analyzed here. Simply put, if much of the diaspora also consists of engineers, as is plausible, then regions abundant in engineers are also likely to be the source of the diaspora, and hence, disproportionately likely to benefit from the diasporic connections. Similarly, if entrepreneurs are more likely to have an engineering background, then regions that produce more engineers will be home to more entrepreneurs.

III. The Indian Software Industry:

According to an early study, Indian software exports were a mere $4 millions in 1980 and rose to $27.7 millions in 1985 (Heeks, 1996). Exports reached $128 millions in 1990. The industry grew very rapidly in the 1990s and exports were over $12 billion in 2003-04.

An important feature of Indian software exports has been the very high human capital intensity, relative to other inputs. Initially, the bulk of the exports consisted of sending software developers to work at the client site in America, on short term assignments. Later, teams of software developers were sent overseas, and only by the mid 1990s, was there significant software activity taking place locally. Initially, physical and communication infrastructure was far less important than commonly believed for the growth of the Indian software exports. Instead, the keys to growth were contacts with potential clients in America and Western Europe, and access to high quality engineers. Indeed, the survey conducted by Arora et al. (2001) referred to earlier found that 57% firms reported manpower shortages as among their most important problems. The next highest problems were employee attrition (44%), market access (42%), and getting visas (33%). Bringing up the rear were physical infrastructure (12%) and lack of government support (10%).

Thus, employment has closely tracked revenues in this industry. The number of professionals was merely 6800 in 1985 and increased more than eight fold in the next five years to 56,000 (Table 1). The growth was at smaller pace in next decade and number of professionals rose to 841,500 in 2003. The number of professionals in the software exports sector has increased more slowly in recent years in comparison to those in the IT enabled services sector (ITES-BPO), from 110,000 in 1999 to 270,000 by 2003.

Table 1: Employment growth in the Indian software industry, in '000s

	1985	1990	1995	1999	2000	2002	2003
Software-export sector				110	162	205	270
Software-domestic sector				17	20	25	28
Software- in-house captive staff				115	178	260	290
ITES-BPO				42	70	180	253
Total	**6.8**	**56**	**140**	**284**	**420**	**661**	**841**

Source: NASSCOM's Strategic Review of 2003, 2004, 2005

Much of what Indian software exports consist of does not require an engineering background, yet software exports from India rely very heavily on engineering graduates. A survey of over 100 Indian software firms in 1997 indicate that 80% of the software professionals employed had engineering degrees, while 12% only had diplomas from private training institutes (Arora et al., 2001). A large fraction of these engineers were not electrical or computer engineers. Instead, these included civil, chemical, textile, and industrial engineers with a 4 year undergraduate degree, though often followed by specialized, non-diploma training in software tools. [9]

There are some important reasons why firms prefer engineers. In interviews conducted in 1997 and 1998, few firms admit to hiring non-engineers, principally due to apprehensions about the signal it might send to potential customers and to other potential hires. The CEO of the fourth largest software firm, interviewed in 1997 said that he hired only engineering graduates from the best possible schools in India. However, this was not because engineering training or knowledge was relevant, but because these students tended to be smart and their backgrounds were useful in signaling quality to potential customers (reported in Arora et al., 2001). Simply put, the undergraduate engineering degree acts a screening device, because of the intense competition for admission to engineering colleges (Spence, 1984).

[9] In recent years, software firms have turned to non-engineers as well, particularly for serving the domestic market, and for IT enabled services.

To be sure, this is not a signaling story alone. Engineering graduates are also exposed to the fundamentals of computers and learn basic programming and sometimes even advanced programming language, reducing need for longer duration training. As well four years of engineering education imparts a set of problem solving skills, methods of thinking logically and learning tools that help quick adaptation to changes in technology, domains and tasks (Arora et al., 2001). Perhaps equally importantly, initially the bulk of software exports consisted of software professionals working on client's site in the US on temporary work permits, or H-1 B visas. US visa requirements meant that it was (remains) easier for engineers to qualify for H-1 B visas.

The software industry also tends to recruit younger workers. Older engineers, settled in other jobs, were less willing to accept jobs with (then) unknown firms, and spend extended periods away from their families.[10] Perhaps equally important was their unwillingness to take on the mundane and tedious tasks that were initially required, the most well known of were the Y2K related projects, where someone had to go through the long lines of code, finding and changing how dates were entered. The upshot of the foregoing discussion is that finding and recruiting engineers was critical for a successful software exporter, and that most of the engineers would be newly graduated, rather than experienced ones.

- **Uneven regional growth of the software industry:**

In the very early period of the 1980s, the software industry was concentrated in Mumbai, the capital of the state of Maharashtra (Heeks, 1996) and also the leading commercial center of the country. As exports grew, the industry spread to other cities and states. Bangalore attracted many multinational companies after Texas Instruments set up its development center in 1985.[11] By 1990 the states of Maharashtra (Bombay), Karnataka (Bangalore), Tamil Nadu (Chennai) and Delhi were the ones with large share of exports and states of Uttar Pradesh (NOIDA), Andhra Pradesh (Hyderabad) and West Bengal (Kolkata) also had software exports, albeit at lower levels.

Many multinational companies (MNCs) set up their subsidiaries after foreign investment norms were liberalized by the federal government in 1991 (Athreye, 2005a). These MNCs' locations were typically in the leading software centers such as Bangalore, Hyderabad, Chennai, Bombay and Delhi (including NOIDA and Gurgaon). Indian software firms typically also had a single Indian location, at least till the early 1990s. Therefore the states which had a

---

[10] Indeed, a recent report indicates that nearly two thirds of the IT workers in India have five or fewer years of experience. http://www.ciol.com/content/services/register/register.asp?fid=1 (accessed 05/25/07)

[11] Source: http://www.ti.com/asia/docs/india/about_tii.html, a*ccessed on 11/06/2005*.

head-start continued to grow rapidly in 1990s. This resulted in very heavy regional concentration of industry. Seven states contributed 95% of total software exports in 2002-03, but only 48% of the country's population, 47% of the net state domestic product (NSDP) and 57% of the industrial production in the country. In the other seven states software exports are growing very rapidly but the absolute size of software exports from these states is still small.

IV. Undergraduate Engineering Education in India

In this section we discuss how India's undergraduate engineering education sector has evolved in past couple of decades. There are three main points to be made. First, there is substantial regional variation in engineering baccalaureate capacity, especially at the birth of the software export industry in the late 1980s. Second, this regional variation is mainly due to differences in private engineering colleges. Finally, the differences in private engineering baccalaureate capacity are are significantly affected by *when* the private colleges were allowed in the state.

Table 2: Engineering baccalaureate capacity in India, 1951-2004

Year	Population in millions	Engineering baccalaureate capacity	Engineering baccalaureate capacity per million of population
1951	361	4788	13
1985	765	45136	59
1995	928	105000	113
2004	1086	439689	405

Source: Our compilations from diverse sources including Ministry of Human Resources Development, Government of India, AICTE, NTMIS.

In India, higher education, particularly technical education, had been provided mostly by the government run institutions, except in last two decades. The majority of the institutions were set up and funded by various state governments. The number of institutions offering undergraduate degree in engineering has increased over the years as also the total intake capacity of these institutions[12]. Table 2 shows the growth in engineering baccalaureate capacity between 1951 and 2004. In 2004 the engineering baccalaureate capacity was 91 times that of 1951. Even accounting for population growth, the engineering baccalaureate capacity per million of population grew thirty fold, from 13 in 1951 to 405 in 2004.

[12] The engineering college capacity of a college/institution is the number of students it can admit in a given academic year in all the disciplines. There are discipline-wise upper limits fixed for each academic year by AICTE. Any increase in the engineering college capacity requires approval of AICTE. The All India Council for Technical Education (AICTE), set up in 1945 as an advisory body, was given statutory status in 1987 through an Act of Parliament. The AICTE grants approval for starting new technical institutions and for introducing new courses or programs.

- ***Regional Variation in engineering baccalaureate capacity***

Table 3a shows the sanctioned engineering baccalaureate capacity by state and year, in hundreds of undergraduate engineers. The last row of the table denotes the year in which private colleges were first permitted (and entered). Table 3 shows large inter-state variation in capacity. In fact share of four states of Andhra Pradesh, Karnataka, Maharashtra and Tamil Nadu was almost 75% in 1990-91[13] as compared to 29% of the population. As other states added capacity, the share of these states has declined, but is still around 63% in 2003. As table 3a shows, the growth in capacity has varied over time and across states. Consider the period from 1990 to 1993. Only three states, Karnataka, Maharashtra and Tamil Nadu were adding capacity. In other states the capacity did not increase perceptibly during these years. Also some states have experienced a sudden jump in the capacity, albeit in different years. These variations are important for our empirical analysis. Finally, states that have significant engineering capacity are those where private colleges enter early, though there are some exceptions. Orissa, for instance, permitted private colleges in 1986 but did not witness significant growth in capacity.

- ***Role of private self-financed colleges:***

The inter-state disparities in engineering baccalaureate capacity are mostly due to differences in the timing and growth of the private sector colleges. Engineering baccalaureate capacity in a state can increase by two ways: either by expanding capacity in existing institutions or by opening new institutions. The new institutions can be in the public sector or the private sector. Much of the actual increase has been through new private colleges.

In 1981, the vast majority of engineering colleges were in the public sector i.e., funded by the federal or state governments and bound by their rule regarding admissions, salary, promotion and tenure. Tuition fees were very low and the vast bulk of the expenses were met from the budgets of the respective state governments, with the exception of the few institutes and colleges directly supported by the central (federal) government. Budget constrained state governments faced severe limits on increasing capacity. Therefore capacity expansion in the public sector has been infrequent, and mostly limited to accommodating new disciplines such as computer science in 1990s and information technology in early 2000s.

---

[13] Their contribution to engineering college capacity was similar even in 1987-88.

Table 3a: Sanctioned engineering baccalaureate capacity in '00s, by state and year.

Year	AP	Delhi	GJ	HR	KA	KL	MH	MP	OA	PN	RJ	TN	UP	WB
1990	58	9	33	5	170	27	192	17	11	5	11	92	31	23
1991	55	10	33	6	180	28	199	19	11	5	11	92	32	23
1992	55	10	34	8	188	29	238	19	11	5	13	94	33	23
1993	55	11	36	8	172	30	256	19	11	11	14	118	33	23
1994	56	10	38	8	193	35	280	19	12	11	14	141	33	24
1995	80	13	44	9	202	45	309	32	12	19	14	185	37	25
1996	86	12	50	9	203	47	333	34	17	19	15	222	44	26
1997	130	13	54	33	238	49	344	48	33	22	15	238	49	26
1998	196	16	64	33	244	51	397	43	45	22	20	273	68	40
1999	241	21	73	47	262	67	429	71	62	22	27	366	85	45
2000	277	23	91	67	282	88	429	102	62	34	50	505	153	52
2001	440	30	106	86	356	113	446	109	88	44	63	655	213	62
2002	624	34	106	98	381	183	470	160	88	86	82	702	231	107
2003	658	35	103	101	389	199	475	194	107	107	115	707	242	107
Year Pvt. College	***'77***	*'99*	*'95*	*'95*	***'57***	*'92*	***'83***	***'86***	***'86***	*'93*	*'98*	***'84***	*'95*	*'96*

AP: Andhra Pradesh, GJ: Gujarat, HR: Haryana, KA: Karnataka, KL: Kerala, MP: Madhya Pradesh, OA: Orissa, PN: Punjab, RJ: Rajasthan, TN: Tamil Nadu, UP: Uttar Pradesh, WB: West Bengal.

Year Pvt. Coll. – The year privately financed colleges first enter in the state

Table 3b: Software Exports by State, 1990 -2003 (in millions of Rupees, 1993-94 constant prices)

Year	**AP**	**DL**	**GJ**	**HA**	**KA**	**KL**	**MH**	**MP**	**OA**	**PN**	**RJ**	**TN**	**UP**	**WB**
1990	127	301	0	0	626	2	1,571	0	0	0	0	374	129	80
1991	158	430	0	0	1,189	6	1,774	0	0	0	0	489	188	94
1992	242	387	0	119	1,595	13	2,521	0	0	0	0	654	248	143
1993	277	986	2	170	2,235	17	3,836	0	3	0	0	1,124	385	251
1994	511	2,849	10	248	3,079	19	4,771	0	4	0	0	1,800	541	319
1995	857	3,459	12	458	4,386	29	6,321	0	9	0	0	2,725	770	410
1996	1,813	5,493	30	745	7,609	62	9,764	0	15	0	0	4,563	1,288	465
1997	2,127	9,458	40	1,442	12,630	196	12,751	0	28	0	0	7,502	1,644	743
1998	4,587	17,643	93	7,763	24,347	374	14,114	106	565	56	26	9,174	7,057	1,411
1999	7,037	26,027	183	6,448	29,158	442	18,703	239	726	99	99	13,171	8,255	2,434
2000	12,103	23,869	641	9,108	48,681	706	26,853	314	1,256	314	188	24,435	21,935	2,946
2001	18,052	14,215	754	17,923	71,786	909	37,866	544	1,545	433	278	36,465	15,451	4,363
2002	22,001	17,121	609	20,377	81,834	958	40,754	604	1,741	406	269	43,529	17,992	7,545
2003	28,152	19,412	782	27,732	107,598	1,212	54,921	693	1,803	1,009	277	44,925	19,689	8,874
Year Pvt. College	***1977***	*1999*	*1995*	*1995*	***1957***	*1992*	***1983***	***1986***	***1986***	*1993*	*1998*	***1984***	*1995*	*1996*

Furthermore, capacity expansion in existing institutions requires approval of All India Council for Technical Education (AICTE). AICTE limits intake capacity of a college. For example, according to the rules in force in 2005, the maximum capacity per discipline was 60 and a college could have maximum of 4 disciplines in first year of its operation. The total capacity could increase by 60 to 300 in the second year and finally to a maximum of 420 in the fourth year. Any increase beyond 420 requires that the institution meet very stringent quality standards, which few do.[14]

The effect of regulation on capacity expansion combined with states' (public sector) constraints in adding capacity means new private colleges have been the main source of growth. Analysis of college-level data between 1981 and 2004 support this conclusion. The number of colleges in the entire country increased from 246 in 1987 to 353 in 1995 and over 1100 in 2003. Eighty percent of new colleges added between 1987 and 1995 were in the private sector and the share of private colleges was even higher at 94 percent for colleges added between 1995 and 2002.

Karnataka was among the first state to permit the private sector in undergraduate engineering education. The first such college opened in Karnataka in 1957[15]. Thereafter one in 1962 and two in 1963 started their operation in the state. Then a large number of private colleges entered, beginning 1979, with nine colleges opening in 1979 and eleven in 1980. The first private college started in 1977 in Andhra Pradesh and in 1983 in Maharashtra after the government introduced policy permitting such institutions to operate. By 1986, only six states had such institutions. Of these, only Madhya Pradesh and Orissa failed to develop leading software clusters. Of the remaining eight states, only Delhi managed to develop a leading software cluster.

As a result, Andhra Pradesh, Karnataka, Maharashtra and Tamil Nadu accounted for almost 75 percent of total engineering baccalaureate capacity in the entire country in 1990. Beginning in 1992, other states began to allow private self-financed institution and by 1999 all fourteen states had allowed private engineering colleges. As a result the share of private colleges has steadily increased over the years, from 62% in 1995 to more than 82% in 2002. In software specific disciplines (principally, electrical and electronic engineering, and computer science and computer engineering), the share is more than 90%.

It is only to be expected that education quality should have suffered greatly during this great expansion in capacity. Many of the new colleges are not up to the task of training

---

[14] Source: AICTE Handbook for Approval Process, 2003-04 and 2004-05.
[15] Manipal Institute of Technology (http://www.manipal.edu/mit/aboutus/overview.htm).

engineers, and their graduates frequently need extended periods of training by employers before they can be put to work. However, as briefly noted earlier, actual engineering skills may be only part of the attraction of engineering graduates, especially in the 1990s. Innate capabilities, a willingness to work hard, and structured problem solving abilities, all of which are likely higher among the graduates of even poor quality engineering colleges, may be more important.[16] As well, in recent years, large Indian firms have undertaken substantial investments in in-house training, in some cases spending 3-4% of revenues on training.

Adding to the problem has been a marked decline in the production of engineering PhDs. The number of engineering PhDs produced fell from 629 in 1991 to 298 in 1996, (AICTE, 1999). This decline in PhD production suggests that while there are strong private incentives to invest in an engineering baccalaureate, these do not extend to investing in research degrees. It also points to the limits of the private sector model for education.

V. Model and empirical specification

This section presents a simple model of a two region economy to motivate our empirical specification. The model shows that if firms are mobile but workers have idiosyncratic preferences for a region, so that workers are imperfectly mobile, then regions with a higher stock of workers will also have greater volume of production.

Many studies have found that a region's growth is influenced by the initial level of human capital. Glaeser et al. (1995) find that human capital level in 1960 influences growth of the cities between 1960 and 1990. Similarly, Simon et al. (2002) found that cities that have higher level of human capital initially grow faster in the long run. Thus initial level of human capital seems to advantage cities and regions, perhaps by attracting knowledge-intensive industries. The regional differences in level of human capital also explain geographic differences in firm formation rates with regions endowed with higher level of human capital having higher firm formation rates (Acs, 2003).

- **A simple model of supply of engineers and size of industry:**

Our objective is to sketch out a simple model to structure our empirical analysis that follows. In our model, both firms and workers choose where to locate. However, whereas firms are profit maximizers and locate in the most profitable location, workers are assumed to have idiosyncratic preferences for their existing location. Workers are homogenous in quality and price taking behavior by firms implies that in any equilibrium where production is not

[16] It is also a tribute to the superior management capability of Indian firms that they were able to use such inexperienced and poorly trained (but bright) young men and women.

concentrated in a single location, wages must be equalized across regions. In such a model it is easy to see that regions with a greater stock of workers will also have more production activity. [17]

Formally, we consider two regions, indexed as 1 and 2. Let N_1 and N_2 be the engineering stock in region 1 and 2 respectively. We assume that the elasticity of supply is zero, i.e., that everybody joins the labor market and is willing to work at prevailing wages, w_1 in region 1 and w_2 in region 2. We further assume that the utility for engineer i from region 1 if she works in region 1 is w_1. The utility in region 2 for i would be $w_2 - C_i$, where C_i is the migration cost (the migration cost includes whatever utility loss there is from moving). Similarly, for engineer j educated in region 2, the utility is w_2 when working in region 2 and $w_1 - C_j$ when working in region 1. We assume that C_i and C_j are all drawn from a distribution F. We do not specify a lower bound for C so that it can take negative values as well. Then, the fraction of workers moving from region 1 to 2 is $F(w_2 - w_1)$, and the fraction moving from 2 to 1 is $F(w_1 - w_2)$. Let $x = w_2 - w_1$. The total labor supply in region 1 is $N_1(1 - F(x)) + N_2 F(x)$, and $N_2(1 - F(-x)) + N_1 F(-x)$ in region 2.

Labor demand: There are M firms, which are price takers. Since the good in question is software for export, we also assume free transport of output. We assume that firms can locate anywhere they want. This is sensible since software is a new industry and most firms are de novo startups. Furthermore, a substantial fraction of the software exports from India are accounted for by American firms and by firms set up by people of Indian ethnicity living in America. All firms have the same production function $Q(L)$. It is immediate that with output price taking firms and labor as the only input into production, we must have $w_2 = w_1$ in equilibrium. So, total labor supply in 1 is $N_1(1 - F(0)) + N_2 F(0)$.

Since the output price is determined in the export market and therefore the same across regions, labor demand and supply will be equilibrated will be through the distribution of firms across regions. If y percent of firms locate in region 1, and if we normalize the price of the

---

[17] Blanchard and Katz (1992) analyze how shocks to labor demand and supply affect short term and long term employment dynamics in the United States. Robak (1982) develops a model of the long run equilibrium with local land and labor markets, with fixed location specific amenities and scarce land, and where firms and workers are mobile. We focus on the long run equilibrium and assume that workers stochastically prefer their existing location.

output p to 1, then each firm employs $m(w)$ workers, given by $Q'(m) = w/p = w$ (Recall that w is same in each region). The labor demand is Mym in region 1 and $M(1-y)m$ in region 2. This yields two equilibrium conditions for the labor market to clear in both regions:

$$Mym(w) = N_1(1-F(0)) + N_2F(0) \quad (1)$$

$$M(1-y)m(w) = N_2(1-F(0)) + N_1F(0) \quad (2)$$

By adding (1) and (2), we get

$$Mm(w) = N_1 + N_2 \quad (3)$$

Equation (3) gives total demand. Substituting for $Mm(w)$ in (1), we get

$$y = \frac{N_1(1-F(0)) + N_2F(0)}{N_1 + N_2} \quad (4)$$

If we let $N_1 = \theta * N_2$, then (4) becomes

$$y = \frac{(1-F(0)) * \theta + F(0)}{1+\theta} \quad (5)$$

In order to understand how share of firms, y respond to changes in capacity imbalance in two regions we differentiate (5) w.r.t. to θ, which equals

$$\frac{dy}{d\theta} = \frac{1-2F(0)}{(1+\theta)^2} \quad (6)$$

This means

$$\frac{dy}{d\theta} > 0 \qquad \text{if } F(0) < 1/2 \quad (7)$$ [18]

- **Empirical Specification**

The simple model developed above suggests that as long as there is some "stickiness" in the labor market, local endowments of human capital will condition the volume of software production in a region. In other words, it suggests that we specify a model with software exports as a function of the stock of engineers. We lack a measure of the stock of engineers in a state over time. However, our measure of engineering baccalaureate capacity is arguably closely

---

[18] This simple model can be easily extended in a number ways. One way is that the idiosyncratic regional preferences may not be symmetric across regions. Regional variations in employment opportunities, career growth prospects and cost of living differences may result in the distribution of C_i being different for each region. Let C_i be distributed with distribution function $F(.)$ in region 1 and $G(.)$ in region 2. In the equilibrium, the wages in two regions are same. Thus, $\frac{dy}{d\theta} = \frac{1-(F(0)+G(0))}{(1+\theta)^2}$, so that $\frac{dy}{d\theta} > 0$ if F(0) + G(0) < 1 (7').

It is obvious from (7') that holding F(0) and G(0) constant, an increase in θ would increase the share of firms y in region 1 provided $G(0) + F(0) < 1$.

related to <u>changes</u> in the stock. Specifically, if there were no mobility of engineers across states, then the growth in the stock of engineers in state *i* would be equal to the (lagged) engineering baccalaureate capacity in the state. Since the available anecdotal evidence suggests that such mobility is in fact small, we use this as a proxy for a change in the stock of engineers. To the extent that there is mobility, we have measurement error. As discussed in greater detail below, we also explore the results of instrumenting for engineering baccalaureate capacity to address biases due to measurement error, as well as problems posted by potential endogeneity. In other words, if S_{it} are software exports in year *t* for state *i*, and K_{it} is the corresponding stock of human capital, we have

$$S_{it} = a_{it} + g_t + \beta K_{it} + \varepsilon_{it}\ . \qquad (8)$$

By taking first differences over time (represented by Δ) we have

$$\Delta S_{it} = \Delta a_{it} + \Delta g_t + \beta \Delta K_{it} + \Delta \varepsilon_{it} \qquad (9)$$

Note that (8) allows for each state to have a different time trajectory for exports, so that the state effect varies by time. For feasible estimation, we assume that $\Delta a_{it} = \alpha_i$ i.e., the change in exports per year (for a given state) does not systematically vary over time. Letting $\Delta g_t = \gamma_t$ yields

$$\Delta S_{it} = \alpha_i + \gamma_t + \beta \Delta K_{it} + \Delta \varepsilon_{it} \qquad (10)$$

In other words, the expected annual increase in software exports is equal to a state fixed effect, a year effect and β times the engineering baccalaureate capacity. This is our benchmark specification. Later, we also report estimates from a related specification where we use as the dependent variable the level of software exports rather than the change. Although the latter specification is not theoretically grounded in our model, it is plausible that with rapidly growing demand, the number of firms may depend not merely on the level of human capital stock but also its growth. Both specifications find support in the data, as discussed below. Also, in both specifications, we exploit the variations in state policy allowing private engineering colleges to develop an instrument for engineering baccalaureate capacity.

VI Data

We obtain data on engineering baccalaureate capacity from the “Annual Technical Manpower Review (ATMR)” reports published by National Technical Manpower Information System NTMIS. These reports are prepared by a state-level nodal center of NTMIS and give details of sanctioned engineering baccalaureate capacity and outturn for all undergraduate technical institutions in the state. The Handbook of Engineering Education, a publication of the

Association of Indian Universities has also been used as a supplementary source.[19] Data on software exports are obtained from the Electronics and Computers Software Export Promotion Council (ESC), which is the apex government trade promotion organization for this sector, for the years 1998-2003. For 1997 and earlier, the ESC does not provide state level export data. Accordingly, we used export revenues by NASSCOM member firms, allocating the export revenues of each firm to the state where its headquarters are located. Till 1995, virtually all firms were located in a single state. Thus, this approximation is a reasonable one. As further described in the data appendix, we verified NASSCOM figures where possible from Dataquest, a trade magazine that has covered the Indian IT industry since 1982, and provides data on sales, exports and employment for the leading firms. For two leading firms, which operated in multiple states, we were able to obtain data on employment by state and allocated export revenues in proportion.

Combining data from two separate sources can lead to problems. For instance, the growth of software exports between 1998 and 1997 yields odd results for some states, particularly for the Delhi, because around this time, firms located in Delhi moved their operations to Gurgaon in the state of Haryana, and Noida, in the state of Uttar Pradesh. Exploratory analysis suggests that these problem are modest, at best. For instance, confining oneself to data from 1998 onwards yields qualitatively similar results. The STPI is another potential source of state level export data. However, for the earlier years only a small fraction of software exports appear to be by companies registered through the STPI. For instance, in Mumbai, in the state of Maharashtra, many of the leading firms were located in SEEPZ, an export promotion zone, and apparently did not report their software exports through STPI. Towards the end of the period, however, exports reported through the STPI are about over 90% of the software exports as calculated by NASSCOM or reported in official Indian statistics.

Carrying out the analysis at the level of the state raises some additional issues. In particular, Delhi is bordered by the states of Haryana and Uttar Pradesh. Software exports from the latter two are concentrated very near their border with Delhi, in Gurgaon and Noida respectively. Since firms can move across the three locations, this results in large jumps and dips in software exports. We chose not to smooth the jumps and dips, principally because doing so does not affect the results.

[19] In a few cases, the data from these reports are inconsistent, typically involving decreases or large increases in capacity or where capacity is markedly inconsistent with the number of graduating engineers. In such cases the other sources have been used to arrive at the acceptable figures of sanctioned intake.

Data on control variables like population, per capita power consumption, industrial output, teledensity, per capita income and number of students graduating from high school (passing the 12th grade) is obtained from various publications and websites of concerned departments of Government of India as detailed in the data appendix. Table 4 shows the descriptive statistics for the variables used in the regressions. Software exports, industrial output and per capita net state domestic product (NSDP) are in constant 1993-94 prices.

The unit of analysis is state even though the industry we are analyzing is mostly located in urban centers in India. To a considerable extent, our hand is forced by the availability of data, since creating measures of the supply of engineers by the relevant metropolitan area, though feasible, is very costly. Moreover, the major software exporting centers are locating at some distance from each other and likely draw upon colleges in the state, at least until the late 1990s. The one exception to this is Delhi, which, as noted already, draws upon Delhi, Haryana, and the western part of Uttar Pradesh.

Table 4: Summary Statistics

	Mean	Std Dev	Min	Max
Software export (Rupees Million, 1993-94 constant prices)	6662	14422	0	107598
Change in Software export (Rupees Million, 1993-94 prices)	1724	4081	-9654	25764
Intake Capacity (number)	11507	14462	525	70660
Outturn (number of graduating engineers)	4923	5731	235	28107
Population ('000s)	57017	36867	9082	183205
Teledensity (no. of telephone lines/100 persons)	3.69	4.78	0.235	41.79
Per Capita Power Consumption (Kilo Watt hours/year)	429	197	148	921
Per Capita Income[a] (Rupees, 1993-94 prices)	7692	3161	3752	17682
Industrial Output[b,c] (Rupees Million, 1993-94 prices)	53608	48285	5739	269843
Electronics Production (Rupees Million, 1993-94 prices)	11426	11455	238	47633
No. of Students Graduating 12th Grade	173213	150213	4521	799916

N = 196. (14 states x 14 years. Some states have missing observations for some variables.)
[a b] Lagged by four years.
[c] Net value addition by all manufacturing units in a given state for each year.

VII Results

We begin with some simple descriptive relationships. Figure 1, which shows the log of software exports by state for three years, points the persistence of export leadership: states which were the early export leaders retain their leadership even after nearly a decade and a half. Figure 2 shows the relationship between software exports and engineering baccalaureate capacity. For all each of the three years, we see a positive correlation between a state's engineering baccalaureate capacity and its software exports (in logs). Finally, figure 3 shows the change in software exports by state in various years against the year in which the state first allowed private engineering colleges. As can be seen, states which allowed private colleges to enter earlier are also those where the software exports have increased the most. Delhi, and its two neighboring states, Haryana and Uttar Pradesh, are outliers in that though they were late in allowing private colleges they have shown rapid growth in software exports. Virtually all the software export growth is due to two suburbs of Delhi, namely Gurgaon in Haryana, and Noida, in Uttar Pradesh. It is possible that this is because Delhi, as the only center in the north of India, may have been able to grow even without a large engineering supply of its own, by tapping engineering graduates from virtually all parts of India except the south and the west.

Overall, these figures suggest that states which allowed early entry by private engineering colleges were also favored early destination of software exporters, and their advantage appears to have persisted even as engineering baccalaureate capacity in other states has rapidly expanded. Finally, figure 4 plots the share of software exports and engineering baccalaureate capacity (not lagged), over multiple years. It shows a marked positive relationship between these two shares.

Figure 1: Software exports by state, 1990, 1995, 2003.

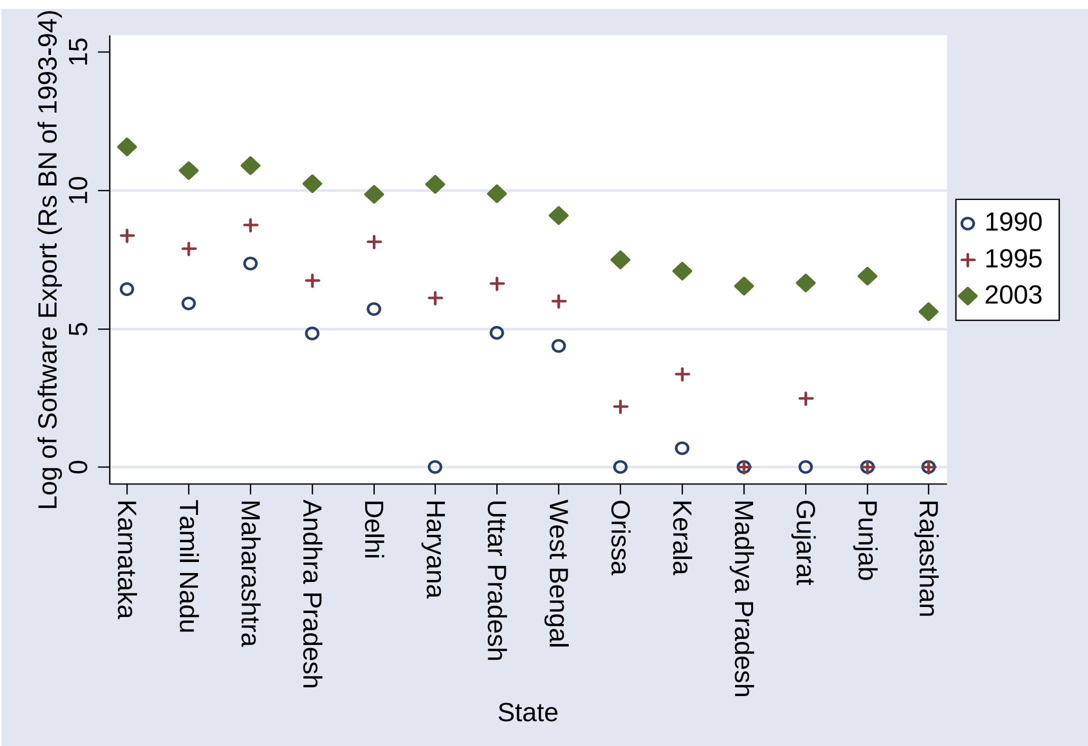
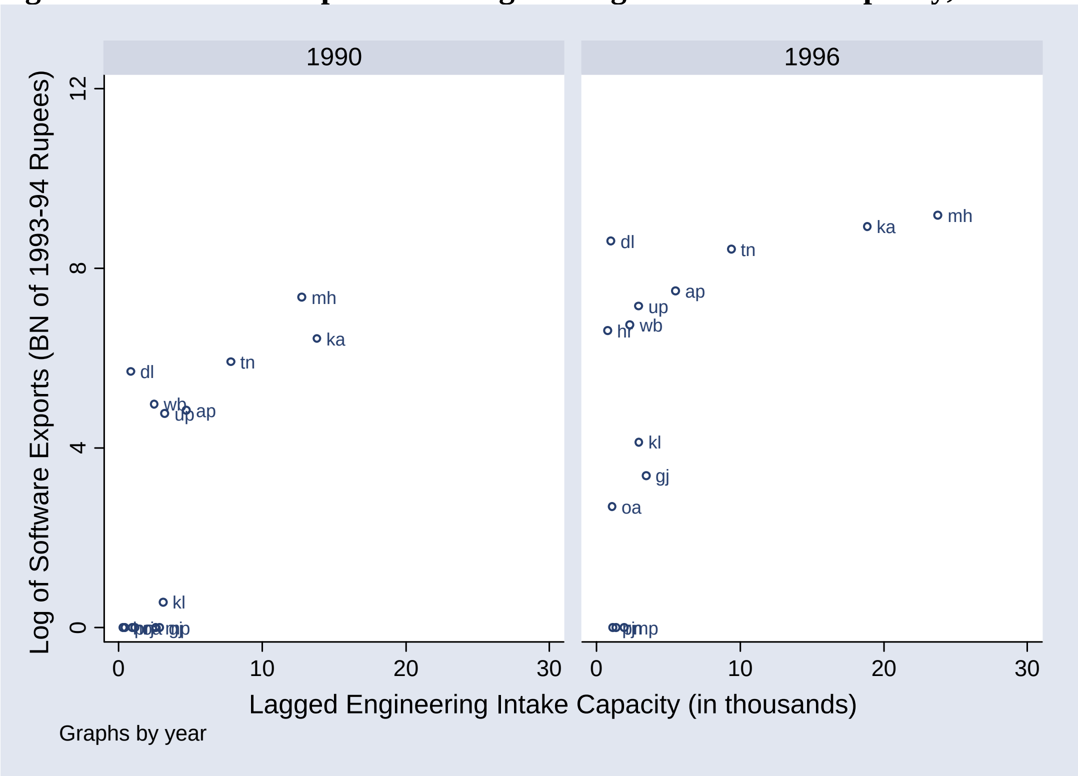
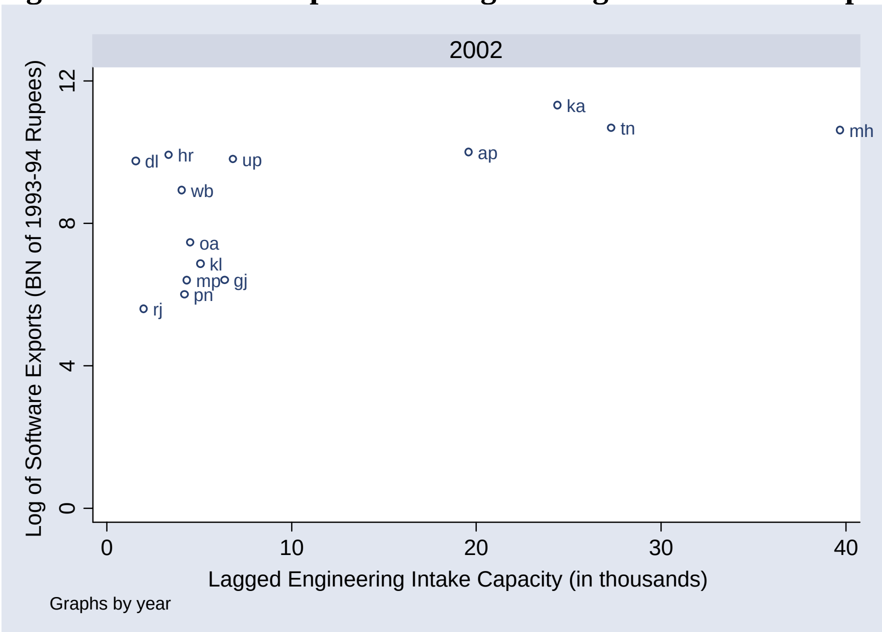


Figure 2A: Software Exports and Engineering baccalaureate capacity, 1990 -1996



Legend: ap: Andhra Pradesh, dl: Delhi, gj: Gujarat, hr: Haryana, mh: Maharashtra, mp: Madhya Pradesh, ka: Karnataka, kl: Kerala, oa: Orissa, pn: Punjab, rj: Rajasthan, tn: Tamil Nadu, up: Uttar Pradesh, wb: West Bengal

Figure 2b: Software exports and engineering baccalaureate capacity, 2002.



Legend: see fig 2a.

Fig.3: Change in software exports 2003- 1990, by year of policy change allowing private engineering colleges.

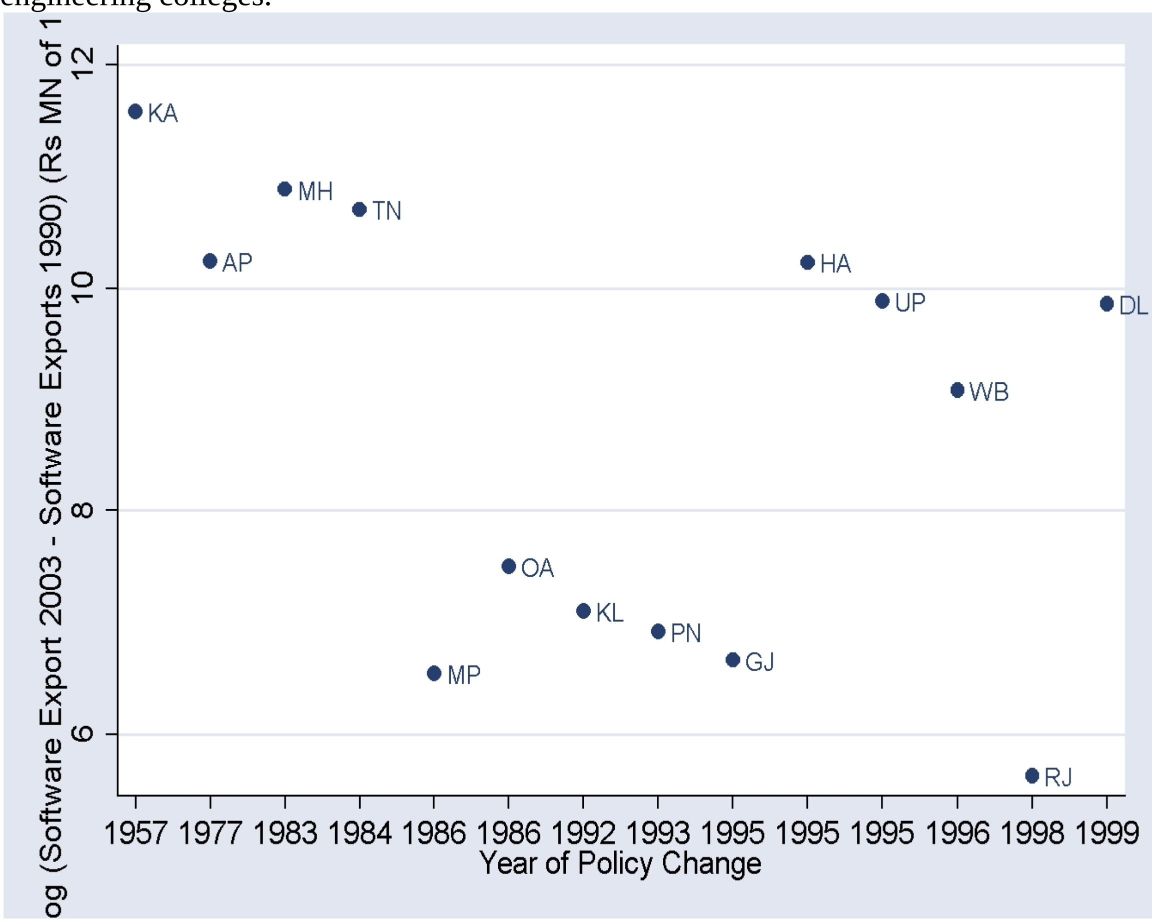
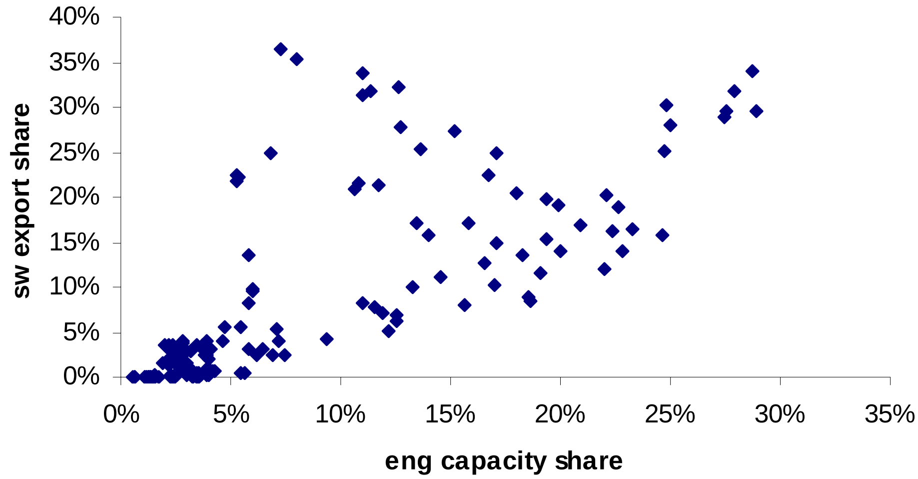


Figure 4: State share of software exports and engineering baccalaureate capacities, 1990-2003.



Notes: From tables 3a and 3b. Delhi, UP and Haryana are combined.

We further explore these patterns through regression analysis. Consider first the long term impact of initial engineering baccalaureate capacity. In table 5 we show the impact of engineering baccalaureate capacity in 1987 on the increase in software exports between 1990 and 2003. It is worth pointing out that the *total* software exports in 1987 were $54 million dollars (Athreye, 2005b) and therefore engineering baccalaureate capacity in a state in 1987 was unlikely to be influenced by software exports industry. As table 5 shows, initial levels of college capacity in the state have a significant and sizable effect upon software exports in 2003, nearly a decade and a half later. The limited number of observations rules out the use of more controls. It is possible, therefore, that this long lasting influence is merely a reflection of unobserved state characteristics. Accordingly, we exploit the within state-variation in capacity over time in table 6, which uses both year and state fixed effects. In addition, we control for per capital income, and teledensity, and use the state's population to control for size effects.

Table 5: Initial state level engineering baccalaureate capacity and software exports

	Software exports 2003 minus Software exports 1990
Eng. Baccalaureate Capacity 1987	5.96 (1.00)
Electronics Production 1990	0.97 (0.50)
Lagged Industrial Output 1987	-0.56 (0.15)
Constant	6096 (4956)
R^2	0.90

No. of obs. 14.
Software exports measured in constant 1993 rupees, millions.

In subsequent analysis, we control for several other factors that might have facilitated growth of the software exports in the state. *Electronics production in 1990* is the size of hardware electronics industry in 1990, before the software industry achieved significant size. We include it in our regression as it has been argued that in the initial years of its growth, the software industry also relied on experienced professionals working in the electronics industry to meet its manpower requirement (Lateef, 1997). This also controls for a variety of other influences. For instance, Klepper (2007) has argued that related industries are more likely to spawn successful firms. However, since firms in many sectors, such as banking, finance and manufacturing are also significant producers of software (primarily for their own use), we control for industrial production as well.

Table 6 reports on the specification implied by equation (10). We lag engineering baccalaureate capacity by four years as it takes four years to complete an undergraduate engineering degree. This makes it unlikely that our effects reflect the feedback effect of software growth, except possibly towards the end of our sample period. Other controls such as electronic production, industrial output, per capita income and teledensity are also lagged, albeit by one year. Further, the standard errors are cluster corrected to account for the non-independence of errors within a state.

Table 6: Annual change in software exports, 1990-2003, (Rs millions, constant 1993-94 prices)

	(1)	(2)
Eng. Baccalaureate Capacity (4 yr lag)	0.34 (0.1)	0.20 (0.07)
Lagged Electronics Production		0.40 (0.24)
Lagged Industrial Output		0.007 (0.023)
Lagged Per Capita Income		-0.55 (0.61)
Population		-0.28 (0.16)
Constant	-371 (1308)	22981 (11914)
State-fixed effects	Yes	Yes
Year-fixed effects	Yes	Yes
R^2	0.49	0.54

Note: Cluster corrected std. errors in parenthesis. No. of obs. 182.

Table 6 shows two specifications, with and without time varying state characteristics such as per-capita income, population and electronics and industrial production. In specification 1, a unit increase in capacity increases exports by Rs 340,000 or about $8,000 per year. To put this in perspective, the average revenue per employee in the software industry in India in the mid 1990s was of the order of $15,000. If one takes into account the less than full capacity utilization, students leaving prior to graduation, employment in industries other than software exports, and migration to other states and overseas, the quantitative impact is highly plausible.

As can be seen in column 2 of table 6, controlling for time varying characteristics reduces the impact of engineering baccalaureate capacity on software exports, but the impact remains both economically and statistically significant. Further, with the exception of electronics production, none of the time varying characteristics added have statistically significant coefficients.

- **Potential endogeneity of engineering baccalaureate capacity**

The identification thus far relies on the fact that the vast bulk of the growth in engineering baccalaureate capacity after 1990 in a state is privately financed, and that differences in the extent of privately financed colleges is overwhelming reason for variation in engineering baccalaureate capacity, both across states and over time. The principal source of variation in the extent of privately financed colleges is when a state permits such colleges -- the earlier the state permitted colleges, the more quickly capacity could increase. It was not until the 1990s that Indian states actively began to compete to attract businesses to locate. Before that, states frequently viewed private business with some suspicion. Though more business friendly states might, prior to the 1990s, have offered tax concessions or regulatory relief, they were unlikely to make significant policy changes in education policy to address business concerns. Moreover, recall that these results control for state fixed effects, which implies that only time variation in the extent to which a state is business friendly would be a source of problems.

Despite this, it is possible that capacity is correlated with unobserved time varying effects that condition software exports from a state. For instance, a growing software industry in a state may create the expectation of growth in future demand for engineers. This will bias our estimate of the coefficient of lagged baccalaureate capacity upwards. On the other hand, it is likely that capacity is an imperfect measure for the change in the state level stock of human capital, and the resulting measurement error would imply an attenuation bias to zero. To probe the robustness of our results to both sets of concerns we present the estimates using an instrument for engineering baccalaureate capacity.

Our instrument is based on the year when neighboring states first allow private engineering colleges to operate. Specifically, we create a dummy variable, “policy”, for each state, which is one if private engineering colleges are operating in that state in that year, and zero otherwise. Our instrument is the average of “policy” for neighboring states. Thus, the key assumption underlying our identification strategy is that though a state’s decision to allow private colleges may respond to anticipated demand for software workers in that state, the decision of neighboring states is independent of the anticipated demand in the reference state. Put differently, we assume that although state governments may respond to local software firms, they are not responsive to software firms in neighboring states. However, states do respond to policy changes in neighboring states. Thus, this is a plausible instrument, though we acknowledge that such instruments have a "reflection" problem (cf Manski, 2000). Thus, we view this as a way of probing the robustness of our results, rather than our benchmark specification.

Table 7a: IV estimates - First-stage results

Dependent Variable: Lagged Eng. Baccalaureate capacity by state

	(1)	(2)
Average neighboring state policy	-6713	-5144
	(2618)	(2429)
Lagged electronics production		0.37
		(0.18)
Lagged industrial output		0.06
		(0.03)
Lagged per capita income		0.1
		(0.82)
Population		-0.24
		(0.11)
Constant	24204	33081
	(4918)	(11098)
State-fixed effects	Yes	Yes
Year-fixed effects	Yes	Yes
F-statistic for instrument	6.33	4.49
R^2	0.89	0.93

Note: Cluster corrected std. errors in parenthesis. No. of obs. 182.

Table 7b: IV estimates. Second Stage Results

	Change in Software Exports (2SLS) **(1)**	Change in Software Exports (2SLS) **(2)**	Software Exports (OLS) **(3)**	Software Exports (2SLS) **(4)**
Eng. Baccalaureate Capacity (4 year lag)	0.62	0.74	1.31	2.02
	(0.36)	(0.50)	(0.27)	(1.95)
Lagged Electronics Production		0.21	1.60	1.35
		(0.23)	(0.80)	(0.56)
Lagged Industrial Output		-0.03	-0.11	-0.16
		(0.05)	(0.08)	(0.20)
Lagged Per Capita Income		-0.67	0.98	0.83
		(0.67)	(1.68)	(1.54)
Population		-0.15	-0.73	-0.56
		(0.14)	(0.56)	(0.35)
Constant	-4773	9397	43376	25642
	(4489)	(11527)	(37591)	(34762)
State-fixed effects	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes
R^2	0.45	0.44	0.75	0.73

Note: Cluster corrected std. errors in parenthesis. No. of obs. 182

Table 7a shows the results of the first-stage regression of college capacity on average neighbor policy, with state and year fixed effects, with and without time varying controls. Though the neighbor policy measure is significant, the F statistic is only around 4.5 with time varying controls, and around 6 without them, after cluster correction. This implies that the instrument is not very powerful. With this caveat, we proceed with the estimation, in part to probe the sensitivity of our results.

Table 7b presents the corresponding estimates where we instrument for lagged engineering baccalaureate capacity using its predicted value, using a two stage least squares procedure. Columns 1 and 2 present results where the dependent variable is change in software exports over the previous year. Note that the estimated coefficient increases three fold as compared to the OLS estimate, suggesting that upward bias was unlikely and that measurement error is more likely. However, the coefficient is imprecisely estimated, possibly because the instrument is weak. We also present analogous results where we use the log of software exports, rather than the annual change in software exports, in columns 3 (OLS results) and 4 (2SLS). The estimated coefficient of capacity in this case also increases upon instrumenting for it, although the increase is not as large. Once again, the estimated coefficient has a large standard error. The other noteworthy point is that lagged electronics production also has a positive and significant impact on the level software exports, but not on the annual change in software exports. This may point to either knowledge spillovers from electronics, or some sort of time varying state characteristics related to IT production. Other time varying controls are statistically insignificant. Specifically, per capita income and industrial output do not play any role in explaining software exports growth.

- **Alternative explanations**

Infrastructure: To further probe the robustness of our findings, we briefly discuss possible alternative explanations for our findings. The first is that states such as Karnataka and Maharashtra were better endowed with telecommunication or physical infrastructure (cf., Srinivasan, 2006). We control for state fixed effects, and also lagged state telecommunication density, lagged state per capita income and lagged industrial production, which should control for variation over time in these effects. Despite this, the ability of the Delhi region to emerge as a significant software exporter despite only belatedly allowing private engineering colleges (and thus having only a small stock of engineering baccalaureate capacity) may point to the importance of an adequate infrastructure. It is likely, however, that this also points to the importance of a commercial infrastructure, including the supply of entrepreneurs.

__Early mover advantage and self-reinforcing effects__: It is popularly believed that software production is marked by significant agglomeration economies. Thus, states which get an early start in software production are more likely to persist as leaders. Arguably, the governments of these states would also be more sympathetic to the need to produce more engineers and the academic entrepreneurs more willing to create private colleges to meet that demand. Such dynamic explanations are not easy to test and in any event, this one has some measure of truth. As figure 1 shows, states that were early software exporters continue to lead in software exports. However, the advantages of an early start do not overwhelm other factors. As can be seen from tables 3a and 3b above, in 1990, Maharashtra had three times the exports of Karnataka, yet by 2003, Karnataka had twice the exports of Maharashtra. Moreover, even if the time series results are driven by increasing returns, if human capital supplies provides the basis for early leadership, an agglomeration economies based explanation need not be inconsistent with a human capital based one.

It is unlikely that the agglomeration economies are due to knowledge spillovers, given the simple nature of Indian software exports. It is possible that a region such as Bangalore, well endowed with research institutes and natural amenities, enjoyed a good reputation with potential entrants, particularly multinational firms. This entry, directly or indirectly, increased demand for labor and also the supply of engineering baccalaureate capacity. Since Karnataka was also the leader in engineer baccalaureate capacity from the very beginning of the software industry, and we lack a direct measure for reputation, we cannot definitively distinguish this explanation from the engineering capacity based one. Insofar as Bangalore's reputation related largely to its supply of engineering talent, this explanation too ultimately supports the human capital story, though not in all the details regarding the role of private engineering colleges.

An under-explored but potentially important source of regional agglomeration is entrepreneurship (Klepper, 2007). Athreye (2005: p 12) estimates that entrepreneurial firms accounted for over a third of the employment and revenue in the Indian software industry in 2001. Among the leading Indian software exporters, more than half were either de novo startups or spawned from other leading software producers (including multinationals). If a state got a head start and was home to successful firms, these firms are likely to spawn other firms, which may reinforce the initial advantage of that state. Insofar as entrepreneurs in the software industry are more likely to be trained as engineers themselves, the location of engineering colleges may be an important source of variation in the supply of entrepreneurship, particularly in the early years of the industry. Therefore, the human capital effects may work through the supply of entrepreneurs.

Diaspora: As noted earlier, the other key factor for export success is contacts with potential clients. Kapur (2002) has pointed to the importance of the Indian diaspora in facilitating such contacts, and Arora, Gambardella and Klepper (2005) provide some evidence of the role of the diaspora in creating firms, including some of the top software exporters. It is possible that the successful states were disproportionately the source of the Indian diaspora. The diaspora explanation is indeed consistent with our results. Anecdotal evidence suggests that a large fraction of the people of Indian ethnicity living in America have engineering undergraduate degrees. If so, state with larger engineering baccalaureate capacity, particularly in the 1980s and early 1990s, were more likely to have produced engineers who emigrated. Thus, though in our discussion thus far we have focused upon the role of engineers as software developers, this is not inconsistent with some fraction of these engineers also forming the diaspora which connected Indian software exporters (in their home states) with their customers in America, or themselves setting up software firms to service American clients.

VIII. Discussion and conclusions:

The importance of human capital -- skilled and creative workers -- to a "high-tech" industry is routinely acknowledged but often public policy discussions tend to focus on more trendy prescriptions such technology parks, venture capital, incubators and university industry centers. Software, perhaps more than any other high-tech industry, relies more intensively upon human capital. Software services, the engine of the Indian software sector, is arguably even more human capital intensive than software products. Thus, few would question the role of human capital stocks in the rise of the Indian software industry. What is less clearly appreciated is that there are significant variations across Indian states in stocks of the relevant human capital, engineers, and that these differences have played an important part in conditioning where the software industry has flourished. Even less well understood are the reasons for this regional disparity in human capital stocks.

We find that these variations exist even after controlling for factors such as how rich or large the state is, and measures of industrial production, electronics production or telecommunication investment. Since engineering education has been controlled and, in the main, provided by state funded colleges, differences in the willingness of states to invest in engineering colleges could, but do not, explain the bulk of the inter-state variation. Instead, it is the role of private engineering colleges which is the key the puzzle. Simply put, states which allowed private engineering colleges to enter early were able to get a head start and, this early advantage has persisted for nearly a decade and a half.

Permitting privately financed colleges helped mitigate the adverse effects of the lack of public investments in higher education. It did not completely ameliorate the problem because, as noted earlier, there has been a marked fall in the production of engineering PhDs, even as baccalaureate capacity has increased. Knowledgeable observers of the Indian software industry, and the leading firms themselves, are increasingly concerned about the divergence, which also points to the limits of relying solely upon private financing for human capital development.

References:

Acs, Z.J. and C. Armington (2004), "The impact of geographic differences in human capital on service formation rates", *Journal of Urban Economics*, Vol. 56, pp. 244-278.

AICTE (2006), "Approved Institutes,http://www.aicte.ernet.in/ApprovedInstitute.htm, accessed on 02/14/2006.

AICTE 1999: 'Reshaping Postgraduate Education and Research in Engineering and Technology', All India Council for Technical Education, New Delhi.

Arora, A., V. Arunachalam, J. Asundi, and R. Fernandes (2001), "The Indian Software Industry", *Research Policy*, Vol. 30 pp. 1267-1287.

Arora, A., and S. Athreye (2002), "The software industry and India's economic development", *Information Economics and Policy,* Vol. 14, pp. 253-273.

Arora, A., Alfonso Gambardella, and Steven Klepper (2005), "Organizational Capabilities and the Rise of Software Industry in the Emerging Economies: Lessons from the History of some US Industries", in A. Arora and A. Gambardella (Eds.), "*From Underdogs to Tigers*", Oxford University Press.

Arora, A. and A. Gambardella (2005), "Bridging the Gap: Conclusions", in A. Arora and Alfonso Gambardella (Eds.), "From Underdogs to Tigers", Oxford University Press.

Arrow, K. (1973), "Higher Education as a Filter", *Journal of Public Economics,* Vol. 2, pp. 193-216.

Athreye, Suma S. (2005a), "The Indian Software Industry", in A. Arora and A. Gambardella (Eds.), "*From Underdogs to Tigers*", Oxford University Press.

Athreye, Suma S. (2005b), "The Indian Software Industry and its Evolving Service Capability", *Industrial and Corporate Change*, Vol. 14, pp. 393-418.

Audretch, D., and M.A. Feldman (1996), "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review*, Vol. 86, pp. 630-640.

Banerjee, A.V. and Esther Duflo (2004), "Growth through the Lens of Development Economics", MIT, Department of Economics.

Berry, C. R., and E.L. Glaeser (2005) "The divergence of human capital levels across cities", *Papers in Regional Science, Vol.* 84, pp. 407–444.

Blanchard, O. J., L. F. Katz; (1992) "Regional Evolutions**"** *Brookings Papers on Economic Activity*, Vol. 1992, No. 1. (1992), pp. 1-75.

Carrington, W.J. et al. (1996), "Migration with Endogenous Moving Costs", *The American Economic Review*, Vol. 86, pp. 909-930.

Creative Infocity Limited, "Infocity Gandhinagar http://www.theinfocity.net/index.htm , accessed on 02/11/2006.

Dossani, R. (2005), "Origins and Growth of the Software Industry in India, Shorenstein APARC, Stanford University.

Ellison, G., E.L. Glaeser, and W. Kerr (2007), "What Causes Industry Agglomeration? Evidence From Coagglomeration Patterns", Working Paper CES 07-13, Center for Economic Studies, Census Bureau, Washington DC.

Feldman, M.P. (2000), "Location and Innovation: The New Economic Geography of Innovation, Spillovers, and Agglomeration", in Clark, G.L., et al. (Eds.) "*The Oxford Handbook of Economic Geography*", Oxford University Press.
Fujita, M., and J. Thisse (1996), "Economics of Agglomeration", *Journal of the Japanese and the International Economies*, Vol. 10, pp. 339-378.
Fujita, M., P. Krugman, and A. Venebles (2001), "*The Spatial Economy: Cities, Regions, and International Trade*" MIT Press.
Glaeser, E.L., J. Scheinkman, and A. Shleifer (1995), "Economic Growth in a Cross-section of Cities", *Journal of Monetary Economics*, Vol. 36, pp. 117-143.
Heeks, Richard (1996), "Indian Software Industry", Sage Publications, New Delhi, India.
Hira, R. (2004), "U.S. immigration regulations and India's information technology industry", *Technology Forecasting & Social Change*, Vol. 71 pp. 837-854.
Johansson, J. and J.M. Quigley (2004), "Agglomeration and Networks in Spatial Economies", *Papers in Regional Science*, Vol. 83 pp. 165-176.
Kapur, Devesh (2002), "The Causes and Consequences of India's IT Boom" *India Review*, Vol. 1, pp. 91-110.
Klepper, S. (2007), "Silicon Valley – A Chip off the Old Detroit Block" unpublished manuscript, Carnegie Mellon University, Pittsburgh.
Lakha, Salim (1994), "The New International Division of Labor and the Indian Computer Software Industry", *Modern Asian Studies*, Vol. 28, pp. 381-408.
Lateef, Asma (1997), "Linking up with global economy: A case study of the Bangalore software industry", working paper, International Labor Organization, New Delhi.
Manski, C. (2000), "Economic Analysis of Social Interactions", *The Journal of Economic Perspectives*, Vol. 14, pp. 115-136.
Maharashtra Industrial Development Corporation's Information Technology Parks, www.maharashtraitparks.com , accessed on 02/11/2006.
National Association of Software and Services Companies (NASSCOM) (2003), "*Strategic Review 2003: The IT Industry in India*".
National Association of Software and Services Companies (NASSCOM) (2004), "*Strategic Review 2004: The IT Industry in India*".
National Association of Software and Services Companies (NASSCOM) (2005), "*Strategic Review 2005: The IT Industry in India*".
Roback, J. (1982). "Wages, Rents, and the Quality of Life**"** *The Journal of Political Economy*, Vol. 90, No. 6. pp. 1257-1278.
Rosenthal, S.S., and W.C. Strange (2004), "Evidence on the Nature and Sources of Agglomeration Economies" in Henderson, V.I. and J.F. Thisse (Eds.), "*Handbook of Regional and Urban Economics", Vol 4*, Elsevier Press, North-Holland.
Saxenian, AnnaLee (2000), "Bangalore: The Silicon Valley of Asia?".
Simon, C.J. and C. Nardinelli (2002), "Human Capital and the Rise of American Cities, 1900-1990", *Regional Science and Urban Economics*, Vol. 32 pp. 59-96.
Simon, C.J. (1998), "Human Capital and Metropolitan Employment Growth", *Journal of Urban Economics*, Vol. 43 pp. 223-243.
Sjaastad, L.A. (1962), "The Cost and Returns of Human Migration", *Journal of Political Economy*, Vol. 70, pp. 80-93.
Software Technology Parks of India, New Delhi, "*First Annual Report*".
Storper, M. and R. Walker (1983), "The Theory of Labor and the Theory of Location", *International Journal of Urban and Regional Research*, Vol. 7, pp. 1-43.
Subramanian, C.R. (1992), "India and the Computer: A Study of Planned Development", Oxford University Press.
Technopark, http://www.technopark.org/index.htm , accessed on 02/11/2006
Tidel Park, http://www.tidelpark.com/index.asp , accessed on 02/11/2006.

Wolter, K. (2004), "High-tech Industry Clustering Rationales: the Case of German Biotechnolgy", in Cooke, P. et al. (Eds.), "*Regional Economies as Knowledge Laboratories*", Edward Elgar.

Appendix

Private Self-financed Institutions:

These institutions are privately founded and operated, with no financial support from the government. The government exercises no control on their day-to-day functioning though there are various regulations by regulatory body AICTE. The creation of new institution, an increase in the capacity of a discipline or the addition of a new discipline in an existing institution requires the approval of the AICTE (and prior to AICTE, approval from the relevant state government). Their principal activity is undergraduate education, and their operations are principally financed from tuition revenues.

Tuition fees were set by state governments. The fee is same throughout state though there are inter-state variations, though the final structure awaits the resolution of court challenges. These institutions are affiliated to universities (universities are set up by an act of state legislature or by an act of parliament), which prescribes syllabi for various disciplines and conducts examinations. The degrees are awarded by the universities.

Appendix B: Data Source:

The main source of data is the "Annual Technical Manpower Review (ATMR)" reports published by National Technical Manpower Information System NTMIS. These reports are prepared by a state-level nodal center of NTMIS and give details of sanctioned engineering baccalaureate capacity and outturn (numbers graduating) for all undergraduate technical institutions in the state. The ATMR has information on sanctioned intake and outturn. The NTMIS publication "Directory of Technical Institutions" has institution level details of intake and outturn. Sometimes the data from these reports are inconsistent. That is the trend in the intake or outturn appear anomalous. In such cases the other sources have been used to arrive at the acceptable figures of sanctioned intake. The Handbook of Engineering Education, a publication of the Association of Indian Universities has also been used as a supplementary source. We have also used printed publications and web-published data of AICTE whenever needed. In certain cases the website of the institution has been useful in providing relevant information.

The data on software exports are obtained from various reports published by NASSCOM and ESC[20]. ESC compiles state-wise software exports data for the post 1997 period. For the earlier period, we rely upon NASSCOM, which publishes the "Indian Software Directory" which have details of software exports, employment and company location. The Dataquest[21] magazine carries a detailed survey of software companies and provides useful information about various software companies, which were used to verify NASSCOM data, and fill in where data on software exports were missing. Our dataset also includes variables like population, per capita power consumption, industrial output[22], teledensity, per capita income and number of students passing 12th grade for each state and for each year. The information on these variables is obtained from various publications and websites of concerned departments of Government of India.

---

[20] ESC (Electronics and Computer Software Export Promotion Council) is India's apex trade promotion organization. ESC is actively engaged in the promotion of India's export of computer software and services, computer hardware, consumer electronics, telecom equipments and cables.

[21] Dataquest covers information technology industry in detail since 1982. They provide details on exports, turnover and employees of software companies.

[22] Industrial output is the net value additions by all manufacturing units in a given state in a given financial year. The data is taken from Annual Survey of Industries conducted by Government of India.

Table A1: Number of engineering baccalaureates by state and year, 1990-2003

Year	AP	Delhi	GJ	HA	KA	KL	MP	MH	OA	PN	RJ	TN	UP	WB
1990	3927	630	2137	235	9015	2325	1940	9575	723	331	1054	5861	2412	1857
1991	4368	846	2164	465	8663	2319	1798	10900	824	307	1109	6147	2412	1959
1992	4385	900	2372	547	9169	2161	2268	12339	805	337	1131	6349	2412	1587
1993	4367	994	2742	533	7665	2246	1823	14323	845	429	1111	6595	2502	2106
1994	4405	847	2852	625	11494	2157	1651	14742	870	522	1265	6669	2502	2304
1995	5610	940	3132	621	11611	2547	2123	15283	851	554	1338	6660	2610	2301
1996	6298	910	3087	683	12182	2441	1849	13772	901	679	1429	7835	2886	2241
1997	5900	1160	3158	662	11977	2795	1647	16812	913	813	1217	9111	2749	2439
1998	5390	1097	3168	657	12036	3001	1763	19516	994	816	1443	11941	3294	2432
1999	7817	1085	3851	1004	12259	3571	2287	20534	1181	679	1469	13452	3323	2518
2000	8102	1103	4723	1120	12526	3877	2158	19706	1498	1365	1445	15524	3552	2644
2001	12171	974	4762	1788	14173	4126	2727	26341	2950	1991	1693	16670	4822	2754
2002	14680	1160	4902	2225	14195	3764	2050	26791	3259	2081	1901	20550	6703	3459
2003	20099	1089	6944	1950	14550	3944	3439	27157	4316	2944	1976	28107	8083	3834

KA: Karnataka, TN: Tamil Nadu, MH: Maharashtra, AP: Andhra Pradesh, DL: Delhi, HA: Haryana, UP: Uttar Pradesh, WB: West Bengal, OA: Orissa, KL: Kerala, MP: Madhya Pradesh, GJ: Gujarat, PN: Punjab, RJ: Rajasthan