Do Perceived Benefits or Costs Drive the Demand for Primary Education in Karnataka?

Nithya Vaduganathan*

June 1, 2005

Department of Economics
Stanford University, Stanford, CA 94309
nithyav@stanford.edu
Advisor: Roger Noll

Abstract

Government expenditure on elementary education in Karnataka has been increasing in recent years to improve the access to and quality of schooling. Despite the progress made towards universalizing education, low enrollment rates and high dropout rates continue to exist, especially at the upper primary level. This paper uses a block level data set from the 2004 Karnataka Child Census to explore the tradeoff between perceived costs and benefits that affect the demand for primary education. By modeling enrollment and completion rates as a function of block, teacher/classroom, and school characteristics, the determinants of expected gain and opportunity costs to education are investigated. The results indicate that while both perceived benefits and costs appear to be important in driving the demand for elementary schooling, benefits appear to be especially important to enrollment, whereas completion seems to be affected by both perceived benefits and costs. These findings indicate the complexities of education reform and the importance of allocating expenditures according to appropriate weights to those factors which best address household concerns and are most valuable to improving the quality of elementary schools.

* I would like to thank Professor Roger Noll for his support, guidance, and insightful comments throughout the writing process. Special thanks to Anjini Kochar and Dr. Geoffrey Rothwell for their advice and ongoing encouragement.

Table of Contents

1. Introduction

The theory that education makes people more productive in the labor market and more able citizens in society is now well established (Hanushek 1986). The positive effects of primary education in developing countries, both from the private and social point of view, are especially apparent. Elementary education, as a building block for national growth and development, is associated with reducing poverty, improving health, and enhancing productivity (Marshall, 2004). However, despite the high returns to education, many primary-aged children in developing countries are not enrolled in school. The prevalence of low enrollment rates and high dropout rates is often attributed to two key factors: high opportunity costs and low school quality (Bedi and Marshall 1999). Both of these affect the household decision to invest in schooling.

Data from a recent block level survey in Karnataka, India offers compelling evidence of this tradeoff between the benefits and costs to education. Notably, approximately 60 percent of children aged 6-14 in Karnataka were not enrolled in school due to lack of access to quality schooling while 30 percent were hindered by labor obligations (Karnataka Child Census, 2004). This thesis will focus on distinguishing the relative importance of perceived benefits and costs in guiding the household decision to enroll and keep children in primary school across the twenty-seven districts in the State of Karnataka.

The Indian government's commitment to the universalization of elementary education through legislation is increasingly evident through resource provision to schools. Sarva Shiksha Abhiyan was the central legislation earmarked in 1986 with the goal of ensuring eight years of elementary schooling for all by the year 2010 (Sinha, 2004). Public expenditures in education as a percentage of GNP in India was 3.4% in 1999 and has been growing since (World Bank Indicators, 1999). Progress has been made in increasing demand for education, especially

amongst the poorest households, through policy measures directed towards reducing costs and increasing benefits from education. Cost reforms include the supply of low cost teaching materials, while measures such as improving school infrastructure and providing hot cooked meals are aimed to enhance the benefits of attending school. Despite these specific implementations, the effectiveness of education policy in India is often ultimately constrained by the pressing need for households to meet immediate food and shelter needs, basic amenities which override the future benefits of going to school.

The decade of the nineties in India has simultaneously seen an increase in school access and literacy rates along with significant dropout rates. Despite several indicators which reflect growth of education in Karnataka, the average primary school dropout rate in 2004 was 11.3 percent (2004 Karnataka Child Census). When comparing these figures to developed countries such as the United States, it is important to note that elementary education in India covers grades 1 through 8; or equivalently, schooling for children aged 6 to 14. With 98% of households having access to an elementary school within one kilometer of residence, the focus on education reform in India is shifting from the issues of universal access to specific school quality concerns (Sarva Shiksha Abhiyan, 2004). The quality of education primarily resides with factors such as school facilities, teacher credentials, and educational resources all of which are concentrated at the school level. This study uses proxies for school quality to gain insight into the demand for elementary education in Karnataka.

1.1 Nature and Purpose of Study

This study uses a modern education production function framework in the estimation of a theoretical model to explain the primary drivers of enrollment and completion of primary school

for children aged 6 to 14 in Karnataka. School enrollment and completion are functions of block, classroom/teacher, and school inputs. The expectation is for children to increase the chance of enrolling in school and staying in school if the benefits to schooling are high and the opportunity costs are low. By conducting a detailed analysis of the effects of a series of explanatory variables on enrollment and completion, this study aims to provide a more complete picture of the demand for primary education in Karnataka.

The majority of the data used in this study is taken from the 2004 Karnataka Child Census, a comprehensive household survey implemented by the Government of Karnataka. All data was collected at the block level and includes information on students, teachers, and schools. The analysis in this paper is contained to government aided public schools across urban and rural areas in each block.

The main hypothesis is that the demand for primary education, as measured by enrollment and completion rates, in Karnataka is influenced primarily by perceived benefits rather than costs. The analytical framework builds on previous studies on school attendance in Honduras and Guatemala (Bedi and Marshall, 1999 and Marshall, 2004).

1.2 Contribution of Study

Karnataka provides an interesting space to study the dynamics of primary school enrollment. The dropout rate at the upper primary level is 31.38 percent despite primary schools existing in all habitations with populations of 200 or more. (Department of Primary and Secondary Education, 2003). The demand for schooling is no longer constrained by the lack of access to schooling but rather affected by a series of other factors which this study aims to specifically delineate. By focusing on the costs and benefits of attaining education, this study

will provide a more well-rounded policy analysis of how government expenditures towards public elementary education can be better allocated to ensure more quality education provision.

This study could be of interest to the general public and research community on two grounds. Theoretically, this research is unique in separating out the determinants of demand for primary education on a comprehensive set of public schools across Karnataka. Additionally, this study, by exploring the relationships between numerous explanatory variables that serve as proxies for the perceived benefits and costs of education, is valuable in guiding real policy implications on how to improve schools across Karnataka.

1.3 Structure of Thesis

This paper is divided into five main sections. The first section reviews the literature on the determinants of demand for primary schooling. The studies cited here are based on research from all over the world. The second section establishes a theoretical framework to explain the economics behind the household decision to invest in elementary education. The model is based on enrollment rates and completion rates being the dependent variables. Some of the limitations of using such a model will also be noted. The third portion concentrates on a description of the data and reasons for the choice of the independent and dependent variables. The next section will include a quantitative account of the regression results and relationships between the variables. Regression coefficients, correlation coefficients, and t-statistics are all reported. The last section will concentrate on the significance of the coefficients and the implications for education policy reform.

2. Literature Review

This literature review will focus on three levels of inputs that drive school attendance: child, household, and school characteristics. This section will evaluate the perceived costs and benefits to education that are associated with each of these input streams using existing empirical literature on school attendance in developing countries. This literature review will also provide an empirical framework for understanding the Karnataka Child Census data presented that is used for analysis in this thesis.

2.1 Child Characteristics

The amount of schooling depends on characteristics of the child such as age and gender. Trends in child traits need to be especially noted in developing countries where the opportunity costs to primary education are high. In examining child effects, it is necessary to use household level data as opposed to aggregated data to minimize endogeneity and avoid capturing individual and community level effects.

Age

Higher opportunity cost for older children suggests that age should be negatively associated with school attendance and achievement. Bedi and Marshall (1999) in their study on primary school attendance in Honduras find that age appears to play a small role in determining achievement, with a marginal effect of between one and three points.

However, Chernichovsky (1985) in his study on school enrollment in rural Botswana finds that the number of children aged 7-14 in the household has a positive effect on the demand for schooling. This runs contrary to the quantity –quality tradeoff. Chernickovsky explains this

finding through the diminishing returns to labor in a household given a certain amount of assets. These decreasing returns will in turn trigger some children being assigned to working in the household or on the farm, while others will be "assigned" to going to school. This observed effect on perceived opportunity costs to education is a combination of age and family size factors.

Gender

The literature on the association between gender and academic achievement shows mixed effects. Duraisamy (2000) shows that enrollment is not equally shared by boys and girls in his household study on the demand for primary schooling in Tamil Nadu. This study finds the gender gap in enrollment to be lowest at the primary level with the ratio of girls to boys at 0.86 and much higher at the secondary level (0.69). Interestingly, at any level, among the children who are enrolled, a larger percentage of girls were progressing on schedule than boys. Staying on schedule signifies lower repetition rates, correct grade level for age, and less dropout. Of those aged 11-15, 65 percent of boys were found to be behind schedule compared to only 54 percent of girls. This research indicates that despite the enrollment gap in Tamil Nadu, the girls who do enter the education process are more likely than their counterparts to stay in school.

An economic perspective on the gender issues in the demand for education is provided by Kuenning and Amin's (2000) study on effective schooling programs in Bangladesh. The study shows that income incentives favor sending girls to school over boys because girls attending schools in Bangladesh receive stipends in the form of lower fees and free books. The lower direct costs to enrolling in schools are a major incentive for parents to send their daughters to school over their sons.

While economic incentives have been shown to trigger an increase in female enrollment, there are cases where cultural barriers have a larger effect in increasing gender gaps in schooling especially in developing countries. One such case is presented by Jamison and Lockheed (1987) in a study on primary schooling in rural Nepal. Only 31 percent of females aged 6-8 are enrolled in school. The authors of this study link female school enrollment in these rural areas to parental attitudes. Conservative parents are less likely to send girls to school and retain them for household chores. Additionally, since most girls leave the household upon marriage, there is less incentive to invest in their education. This study shows how it is difficult to capture such trends and effects in a purely quantitative manner.

The demographic characteristics of age and sex are just two variables that attempt to show the opportunity costs associated with attending schools. The different trends that were observed in each population studied indicate that these are not universal patterns but rather specific to a sample at hand. Since opportunity costs are perceived and vary drastically between countries and households, it is difficult to get consistent and conclusive results.

2.2 Household Characteristics

Family Income

Another debated area in the literature has been the role of socioeconomic background and other household characteristics in driving the demand for schooling. In their study on enrollment patterns in Honduras with days of school attended as the outcome variable, Bedi and Marshall (1999) find that children from higher income households attend between five to nine more days of school than those from lower-income families. Families with higher income are likely to have

fewer credit constraints and are able to send their children to school more regularly. The direct costs of schooling such as fees and books may be more prohibitive to poor parents. Income should only be a significant variable if families are credit constrained.

Parental education level

Besides the income level, another major household factor that affects the likelihood of children attending school is the level of parental education, especially the head of the household. Chernichovsky (1985) finds, amongst the households in rural Botswana, the education of the person heading the household to be the most significant factor in determining child schooling. For families where the household head had formal education for 4 to 11 years, 58 percent of the children aged 7 to 18 were enrolled compared to a 28 percent enrollment rate in households where the average education of the head was only 0 to 3 years. A possible policy implication from this study is that education expenditures can be targeted towards adult education in areas where parents and elders have low levels of training and schooling.

Household location

Government subsidies have different effects on enrollment in urban and rural areas. In urban areas, Birdsall (1985) finds a substitution effect with expenditures increasing schooling access and quality, but at the expense of replacing parental support. In contrast, the elasticity of demand is much higher in rural areas where rising expenditures typically coexist with more extensive parental involvement. Rural areas probably have a higher opportunity costs, if children can provide useful labor on the farms, and higher direct costs. Schools are likely to be more scattered in rural areas and thus transportation costs to school can be higher. These

differing elasticities of demand for education indicate that governments can allocate budgets differently to better reflect the concerns in urban versus rural areas.

This section shows some of the ways in which family background and socioeconomic factors effect the demand for education. While the studies mentioned above indicate that socioeconomic status does matter, other studies shows that socioeconomic status seems to have less explanatory power as a determinant for school attendance and achievement in developing countries than for developed nations (Lastra. 2001). Heyneman and Loxley (1983) interpret these results to mean that school factors may be stronger in societies with fewer educational resources and low levels of literacy. So in these areas, background factors may play a less important role to school level inputs.

The Coleman Report is possibly one of the most well known pieces arguing that school quality based on inputs makes no difference to school outcomes. This study focused on racial differences in schooling in the United States in 1966 and concluded that achievement was not easily explained by simple measures such as school resources like expenditures on teachers or classrooms. In this study, family background and other characteristics of students (age, gender) in school seemed to be much more important as explanatory factors in the demand for education. The Coleman Report has triggered lively debate in this field in order to distinguish the importance of school level factors. The central question posed in this thesis seeks to determine the degree to which school level inputs affect the demand for primary education in Karnataka. The next section presents some of the literature surrounding different school level inputs.

2.3 School Characteristics

Education is a quasi-public good where both social and private benefits are important, but the optimum mix of public and private financing of education is difficult to specify. In a compilation of data on public expenditure on education in India, Shariff and Ghosh (2000) find that the annual growth of expenditure on education deteriorated from 10.2 percent of GNP between 1980-81 and 1990-91 to a level of 3.4 percent in 1995-96. Public financing of education is required to attain the social externalities of schooling. Even though schools are not held directly accountable for what happens to students once they exit the system, responsible governments will allocate budgets to get the most cost effective achievement gains and make-up for the public good nature of education. The following are some of the ways in which the school can add some value to a student's knowledge and skill base.

Teacher inputs

Spending on teaching staff is one of the key components of school expenditures. Budget for instructional staff can determine how many and what kind of teachers are available for instruction at a school. Bedi and Marshall (1999) discovered that, in Honduras, teacher quality, measured by the disposition toward active engagement and participation in the classroom, is the most important determinant of academic achievement. Furthermore, a cost-effectiveness analysis indicated that investing in teaching quality was the most efficient use of financial resources. The effect of teacher experience and qualifications are found to be statistically significant across grades, but seem to have a more profound effect for higher grades.

Jamison and Lockheed (1987) find that in Nepal, the poor quality of teachers is a principle reason that children dropout. Many teachers lack certification and only one-third have

training beyond the tenth grade. These less qualified teachers were also using more traditional teaching methods in classrooms with minimal engagement of the student. Students were expected to read aloud from their textbooks and respond to teachers' questions in unison. There was no emphasis placed on encouraging students to think for themselves or question the teacher.

While teacher quality is related to experience and qualifications, other studies point out that teacher quality and effectiveness are limited by the availability of books and physical facilities in a school. Rampal (2004) finds that children in West Bengal are constrained in learning because teachers have minimal materials and supplies. Additionally, there are a number of single teacher schools where one teacher leads multiple grades simultaneously. This study finds academic achievement to be negatively associated with multi-grade classrooms for children in the early stages of elementary school.

Pritchett and Filmer (1999) provide a deeper analysis on the education production function and how it is influenced by teaching materials. They find that the marginal product per dollar of inputs such as instructional materials that are not directly valued by teachers contributes approximately 10 to 100 times the value of inputs that are directly valued by teachers, such as wages. This suggests that high teacher salaries do not necessarily lead to higher quality schools.

One of the most noted studies on class size effects is Krueger's (1999) study in Tennessee. Known as Project Star, this experiment analyzed data on 11600 students and their teachers who were randomly assigned to differed sized classes from kindergarten through third grade. The main conclusions were that performance on standardized tests increased by four percentile points in the first year that students attended small classes and only by one percent subsequently. Krueger explains this observation because attending smaller classes at the entry level in primary schools may provide a one time socialization effect which permanently raises

student achievement therein. However, in developing countries, bigger class sizes coupled with fewer teaching materials may have a larger negative effect on achievement than in developed countries.

Classroom level factors

A key concern for policy makers in developing countries is the cost effectiveness of education reforms. Where will a marginal dollar increase in expenditure contribute most to academic achievement? Glewwe and Jacoby (1994) look to answer this question based on a study population in Ghana. The results show that the indirect effects of improving school quality, through better infrastructure for example, were strongly linked with achievement and attendance. Repairing leaky classrooms was found to be the most cost-effective investment and policy decision in Ghana. Children who attended schools with leaking classrooms did significantly worse on reading and math tests.

School organization and management

Besides teacher inputs and classroom level factors, there are other inputs into the education system that affect the school on the whole and are also important in guiding the demand for education. Purkey and Smith (1983) point out that these organizational variables such as the school culture, site management, instructional leadership, and parental involvement should all be considered in administrative decision making as important contributors to the educational experience that students receive. Explicit attention should be given to program implementation and evaluation to integrate the different inputs such as teaching with school infrastructure to provide a wholesome learning experience.

Decentralization of schooling is becoming an increasingly popular strategy for remedying the problems of accountability with education systems in developing countries. Decentralization is the process by which power is shared or filtered down from the state to the regional or local level. It is about building institutional capacity so that the community is able to set goals, anticipate needs, and make informed decisions with greater responsiveness to the citizens of a particular district or community. Since parents have an important stake in their children's education and ultimately make the decision of whether or not to enroll their child in school, it is crucial that they have a role in this larger discourse on the manner of education provision at the local level.

3. Methodology

The economics of education is dominated by the human capital approach which links education to the labor market. Education is noted as a primary determinant of wages and is included in most estimates of earnings and the demand for learning and training (Belfield, 2000). The economists' role, in developing demand theory, involves understanding the process by which education changes behavior. That is, how can the demand for education be modeled? It is important to note that education and learning are not always easily quantifiable and do not operate in markets with complete certainty about future events. Thus, there are limitations in the application of predominantly quantitative models and methodologies.

Most studies on the demand for elementary education assume that the parent determines the likelihood of school attendance for their children on the basis of expected benefits and costs to attending school (Bedi and Marshall, 2002). Marshall (2004) also argues that in many developing countries, the absence of enforceable laws mandating school attendance for a

particular age group gives the households flexibility in the school attendance decision making process, thereby involving the parents. The economic contribution of children to the household in developing countries and the opportunity cost linked with school attendance may be substantial. Parents essentially face a tradeoff between household consumption now and children's expected future income because schooling is not free to the household. There are direct costs such as books, uniforms, and fees that will reduce consumption in other areas. Attendance will also decrease if parents perceive the return associated with education to be lower than the child's expected economic contribution to the family. Low quality teaching, poor school infrastructure, and limited teaching materials are some factors that may affect the benefits that parents consider when enrolling their children in schools.

Cost-Benefit Analysis

A basic research method in the economics of education is a cost benefit analysis that compares the costs of undertaking an educational program with the benefits (Levin, 1995). While the cost benefit framework is traditionally applied to determine whether a policy is worth undertaking, the fundamental rule of undertaking a program where the benefits exceed the costs apply to the parental decision of school enrollment as well. For schools, the costs would be associated with any inputs (such as expenditures on textbooks, fees, transportation, etc) necessary for education and the longer term benefits would come in the form of outputs such as improved productivity levels.

Using this approach to value education is difficult because the analysis would require pinpointing costs, including opportunity costs, and enumerating benefits in monetary terms. For example, benefits such as improved health and increased societal awareness are hard to quantify

and can easily be omitted from the cost benefit calculation (Belfield, 2000). The length of modeling can also be a limiting factor. Ideally, the costs and benefits should be projected across a range of years to get a more accurate valuation today. In reality, most parents can only foresee immediate costs and benefits over the short term (one to two years) when determining whether or not to enroll their child. Since each year of education adds a different amount of value to the student, assuming a single year is representative of the education cycle is limiting. Another key problem in estimating the benefits of education is selectivity bias where students are more likely to enroll in classes if they expect to succeed and be effective (Belfield, 2000).

Despite these shortfalls, the cost-benefit analysis offers some key insight into the decision to invest in education. Policy makers often use a cost benefit framework when evaluating the potential for large scale reforms such as reducing class size or increasing teacher capacity in increasing the attractiveness of education.

Education Production Function

The education production function describes how education can generate a particular outcome given a series of inputs. The production function model can be tested via OLS regression analysis to look at whether or not increasing resources will result in superior educational outcomes.

Most school effectiveness studies have used test score on mathematics and reading as their main measures of achievement (Bliss, 1999). The rationale behind using this measure is that it has been found that the mastery of fundamental math and reading skills increases the probability that students will stay in school and ultimately join the labor market in a productive manner. For developing countries, these scores are often a criterion to measure basic cognitive

skills that contribute to the building blocks of development. However, the main limitations of these studies are that since schools produce a number of outcomes, it is difficult to delineate a single outcome as a representative measure of output. Math and reading test scores do not necessarily account for the social, artistic, and behavioral outcomes that schools contribute towards.

Glewwe and Jacoby (1994) find that regressing achievement test scores on school characteristics in Ghana, using ordinary least square regression, to be limiting since public school children are not restricted to attending a particular school. When parents have more choice among schools and choose schools based on school and district level attributes become endogenous to the production function. This data set was unique in that it contained information on the characteristics of the schools students attended and of the alternatives that were available to parents.

For the education production function, Carnoy (2001) finds that school attendance is a good proxy for school quality. Field research supports that parents are more likely to send their children to school when the schools are of higher quality (Hanushek and Lavy, 1994). Parents consider quality with school traits such as facilities and infrastructure along with classrooms specific variables, such as teacher quality, when making their decisions. In countries where the opportunity costs of education are high, these perceived benefits on school quality are important in determining the demand for schooling.

4. Econometric Model

In order to explore the demand for primary education in Karnataka, the following econometric model from Bedi and Marshall's (1999) function of school attendance as a function

of the expected gains from schooling and the value of the child's time will be adapted in this thesis. School attendance (the number of days attended during the school year) is specified as,

$$A_i = \alpha + \beta_H H_i + \beta_V V_i + \beta_F F_i + \varepsilon_i$$

where H_i is the measure of the benefit or returns to schooling, V_i is a vector of variables capturing the opportunity cost of schooling, and F_i is a vector of household characteristics. The error term ε_i measures the unobserved factors such as student effort, healthy, and innate ability and is assumed to be normally distributed with zero mean and positive variance.

To answer the question of what drives demand for primary education it is important to look at what drives students to not only enter school but also remain in school. Of particular interest is the completion rate at the elementary school level. The parental decision to invest in education and continue that investment is reflective of relative opportunities that are available at different points in time. Due to the lack of information on school attendance records and the importance of getting a perception of the time horizon of investment in education, Bedi and Marshall's (2002) theoretical model is adapted in this paper to account for these issues.

The use of a cross-Karnataka data set in this study has constrained the unit of analysis. Instead of using household level data as many of the previous studies cited in the literature review have done, this paper uses a block level data set. A block is a cluster of villages. The decision maker is now the "representative average consumer" per block rather than the individual. Additionally, rather than using the vector F for household characteristics from Bedi and Marshall's model, this study uses block characteristics. Since the existing studies have already done an in-depth analysis of child (age and gender) and family (income and education level) characteristics in both developed and developing countries, this thesis will primarily focus on the educational inputs to schools, classrooms, and teachers in its analysis.

The model used for analysis in this study seeks to establish relationships between measures of perceived opportunity costs and benefits and the demand for education. The following basic linear probably models are used to start analysis:

$$\text{Enrollment} = \beta_0 + \beta_1 [B^*] + \beta_2 [C^*] + \beta_3 [S^*] + \varepsilon_i$$

where B* stands for block characteristics, C* is classroom and teacher inputs, S* represents school variables, and ε_i is the error term. The same variables are regressed on completion rate.

Benefits are represented in this model by the expected human capital gains from educational inputs (teacher/classroom factors) and school inputs and facilities that often signal quality of schooling to parents as an added perceived benefit. Educational inputs are characterized by the percentage of trained teachers and the student teacher ratio. School inputs are represented by the presence of a boundary around the school, availability of a toilet facility, and access to a kitchen for provision of lunch through the mid-day lunch program.

There are limitations to this model. First, the use of enrollment and completion rates as dependant variables poses a challenge because the coefficients from a linear model cannot be constrained. Since probabilities would be truncated above one and below zero, there is room for biases in the estimates of the standard errors on the coefficients. These limitations can be addressed by transforming the dependent variables into odds ratios. This will minimize the errors associated with truncation. The next section includes a detailed explanation for the choice and definition of the dependent variables.

Another possible limitation in estimating educational production functions is the endogeneity between the school quality variables. For example, if parents move in response to school quality or districts provide more resources to compensate for low student achievement or parents move in response to school quality, then the effects of school inputs may be biased and

difficult to segregate. In their study on school enrollment in Honduras, Bedi and Marshall (2002) find endogeneity due to parental migration quite unlikely because ninety percent of the individuals in the household survey indicated that they migrated to find new work rather than for family reasons.

Even if parental migration is not a major concern, the block level nature of our data lends to the possibility for common unobservable factors working at the block level to affect the demand for education. These unobservable differences can lead to correlations in the error terms between educational blocks. These biases can in turn affect the standard errors in least squares estimations. Thus, we should be cautious and aware of the limitations in the application of this theoretical model in practice. To minimize the biases associated with endogeneity, we include district level dummies on one set of specifications in the analysis section.

5. Data Description

The data used in this study is taken from the 2004 Government of Karnataka Child Census, a comprehensive block level survey implemented by the Department of Education in Karnataka in conjunction with Sarva Shiksha Abhiyan and the Azim Premji Foundation. This survey was obtained from the Sarva Shiksha Abhiyan office in Bangalore, Karnataka. Data has been collected on the child population aged six to fourteen in each of the blocks across the twenty seven districts of the state. The data was collected primarily to align available information on the inputs and outputs of the education system across the blocks so that policy makers could better monitor the performance of the education system and respond accordingly. The majority of the questions were directed to get a quantitative evaluation of the present state of the education system. This data has been used by the Government of Karnataka as a baseline

assessment survey to support current state government level interventions and expenditures at the block level. In this study, the data will be used to estimate and assess the block level demand function for elementary education.

The main contributors to this data set have a vested interest in education reform in the State of Karnataka. The Department of Education in Karnataka is a large body that covers approximately 50,000 primary schools from standards I to VII across the State. Sarva Shiksha Abhiyan (SSA) is the Government of India's commitment to the universalization of elementary education as mandated by the 86th amendment to the Constitution of India (Sarva Shiksha Abhiyan, 2001). This amendment guarantees access to primary schooling for all in the 6-14 age group within one kilometer of household residence. SSA works with the government to provide better quality schooling and continue to open new schools to address the access ides.. The third partner, the Azim Premji Foundation, is a non-profit organization based in Bangalore, that is working with the State government to implement a standardized Management Information System (MIS) for the Education Department as a reporting tool (Azim Premji Foundation, 2005). The goal is to make the district level data up-to-date and easily accessible to the public.

The surveys were designed in a way to minimize potential biases and maximize comparability across blocks. Even though block data is limiting in assessing age and gender, it is a reasonable and appropriate unit of analysis because most decision making within States is at that level. According to Jain (2004), the education system in India really needs quality control in relation to a technique of monitoring. The intangible concept of education reform needs to be broken down into small measurable tasks that can be planned, budgeted, implemented, and monitored. The level of detail in this baseline survey is unique to education data collection in

India. Information on disabled children, number of toilets per school, and expenditures to repair classrooms is valuable when assessing reforms.

Despite the concerted efforts of three reputed organizations in collected this household survey, it is important to note some of the potential biases and limitations in the reliability of the data. First, the school level data on state of infrastructure and teacher performance is based on school records for which there are reasons to question the accuracy and likelihood of being up-to-date. Jhingran (2004) attributes the lack of transparency in program implementation in India due to the lack of detailed information at the school level. With minimal school supervision or norms of accountability, few schools keep regular attendance records of students and teachers or documentation of expenditures. Second, even though this is a comprehensive data set with urban and rural blocks, it is impossible to include all public schools across the state. The greater distance between villages in rural areas may also add to the difficulty in attaining a comprehensive set. Finally, non-response on some data points from districts poses an additional issue of concern. These are just a few limitations among others that should be kept in mind but may not have significant impact on the results and analysis.

Variable selection

This thesis uses the Government of Karnataka's Child Census data in a new way by placing the variables in the context of a demand function for elementary education. The process of selecting the dependent and independent variables for analysis was difficult and required awareness of the literature in this field along with consciousness of the biases and potential limitations of the data. To answer my central question of whether primary education in Karnataka is driven by perceived benefits or costs, I had to first figure out a way to divide up the

data into streams of benefits and costs. Using the theoretical delineation presented in methodology section and constraints of my data set, I divided my data into three sources of inputs: block characteristics, teacher/classroom, and school level inputs.

5.1 Dependent Variables

School attendance is a good dependent measure because it spans an entire year of schooling. Due to lack of information on school attendance, I had to find a substitute variable or set of variables that would allow me to observe the nature of the demand for primary education in Karnataka. The three main variables that were considered as possibilities for the dependent variable(s) in this study included transition rate from lower primary to upper primary, enrollment rate, and completion rate. In India, the primary school system is divided into lower and upper primary. The transition rate is useful to understand what percentage of students pass grade 5 and continue on to grade 6, however it is limiting from a time horizon perspective in understanding the overall demand for primary education. Since education is a long term investment that has a range of long term benefits, a variable with a broader time horizon would better capture the perceptions of the returns to education.

Enrollment and completion rates were chosen as the basis for the dependent variables in this study because they better account for these trends in returns to education. Enrollment rate is defined as the number of students aged 6 to 14 who are enrolled in schools relative the population of children in that age group. Completion rate is defined as the number of students who complete primary schooling out of the total number enrolled. Due to the correlation between enrollment and completion in that only students who enroll in the first place can complete school, it is important to classify the completion variable in these terms.

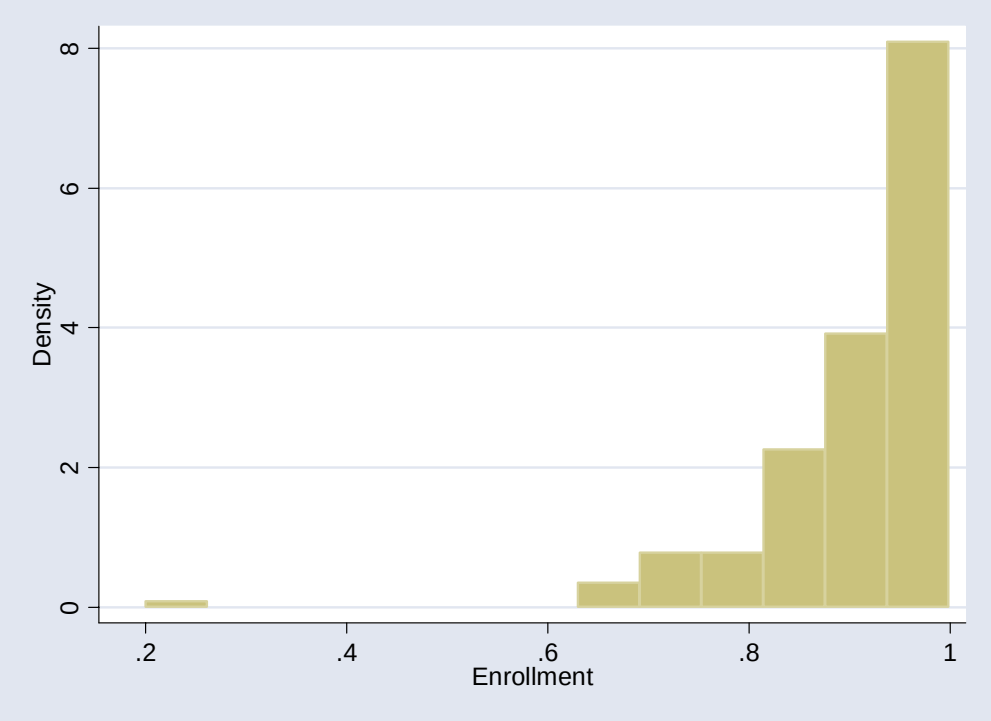
Theoretically, the unit of analysis is the household's decision about a child's education. The probability of enrollment and completion for a child is a function of household and school characteristics. However, since the data set does not contain household level observations, this study uses a "representative consumer" model where block level decisions are driven block circumstances. The representative probability (r) is the average probability of enrollment and completion. Since the probability of enrollment and completion is the same average level for all the households in a block, the total enrollment (E) is the representative probability times the child population (CP).

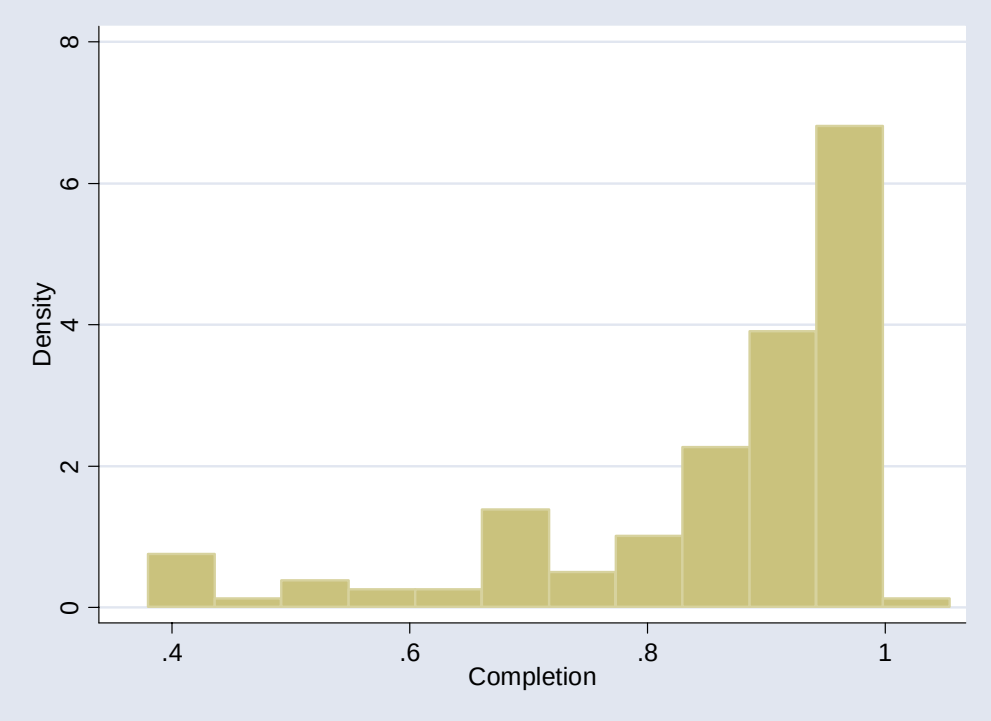
The representative probability of enrollment is a function of a number of inputs and is given by,

$$r = E / CP = f(B^*, C^*, S^*)$$

where B*, C*, and S* stand for block, classroom/teacher, and school level inputs respectively and are indicators of the average values in the block. However since E / CP and C / CP (with C being completion) are ratios, they must fall in the range between zero and one.

The decision to use an odds ratio in measuring enrollment and completion was based on a number of considerations. Histograms of the enrollment and completion rates, as shown below, illustrate that both variables are heavily left skewed. A significant percentage of the blocks in this study have enrollment and completion rates between 90 and 100 percent.





Since the dependent variables lack a normal distribution and are truncated at zero and one as a percentage, using the variables in a simple OLS regression model is limiting and not

necessarily the best fit for prediction. With this data set, there are fewer observed points in the lower percentages that are closer to zero, so the concerns of truncation are primarily concentrated on the upper end. The truncation of values over one in a standard OLS regression leads to biases in the estimation of the variance on the standard errors of the coefficients.

The odds ratio is one way to transform the dependent variables into something without a truncated range. Odds is the ratio of the probability that something is true divided by the probability that it is not. So in this case,

Enrollment odds = % enrolled / (1 - % enrolled) and similarly for completion odds. The odds ratio for enrollment has to do with the effect of moving from being not enrolled to being enrolled. For the purposes of this study, the odds ratio is used to estimate how strongly the independent variables are associated with the outcome of interest.

Due to the nature of data points being concentrated at the tails or ends and the minimal likelihood of the odds ratio being strictly linear, a logit function was used on the odds ratios. Logits contain the exact same information as the odds ratios (they measure the strength of the relationship between variables) but can be compared more easily because they are symmetrical.

Thus the estimating equation becomes:

$$\log [r/(1-r)] = \log A + a[\log B^* - \log (B^*+1)] + b[\log C^* - \log(C^*+1)] + c[\log S^* - \log(S^*+1)]$$

where A is the constant term and a, b, and c are the estimated coefficients. The independent variables must be the different between logX and log(X+1) to prevent the dependent variable from exceeding one.

Logits do not make immediate intuitive sense because we are not used to thinking in terms of odds ratios and natural logs. A positive logit means that the independent variable has the effect of increasing the odds of the dependent variable having a given value. A negative logit

means the independent variable has the effect of decreasing the odds or likelihood that the dependent variables have a stated value. If an independent variable has a logit of .40, this means that the logged odds of the dependent variable change by a factor of 40 percent. For the purposes of this study, we will focus on looking at the significance of the dependent variables in the regression specifications and report the logit coefficients.

The use of two dependent variables in this econometric model adds value to modeling the demand function for primary schooling in Karnataka. The literature suggests that benefits strongly drive the demand for education at the elementary level. The central question posed in this study is to see if that holds true based on this data set from Karnataka and using enrollment and completion information provides us with an idea of whether the factors that induce a parent to enroll their child in schooling initially hold throughout the length of the child's education or if those reasons change over time.

5.2 Independent Variables

The independent variables in this study are classified into block, teacher/classroom, and school characteristics. These were chosen based on what has been used as common determinants of the demand for education in previous studies and what was available in this particular data set. Child-specific variables such as age and sex of the child were left out because they were not available at the household level. They show limited variation at the district level. Instead of child and family, the aggregated data set uses block characteristics of population per village and adult literacy levels. These regressors were included to capture economies of scale effects in educational provision.

Educational inputs are shown through the training of teachers and student teacher ratio. A trained teacher is one who has met NCTE (National Council of Teachers of English) guidelines. For elementary school teachers, this means that they must have gone through two years of training through an education program and an internship at a school for thirty days. This variable is measured as the percent of teachers who are trained. While there have been extensive studies that show that class size has a small effect on school quality, by including this variable we can see if this particular data set confirms the trend or if developing countries actually rely more heavily on having a certain class size. Child population per school and child population per school squared are included as instrumental variables that can stand as proxy\ies for the student-teacher measure. The reasons for choosing these proxies are explained in the next section.

Finally, a set of school inputs are included to gauge parental perception of infrastructure and facilities as added benefits to enrolling children in schools. School lunch programs, also known as Mid-day meals in India, involve having a kitchen facility and appointed cooks to provide daily lunch service for all the children.. The scheme of providing hot cooked meals, according to Karnataka's Sarva Shiksha Abhiyan, is done with the objective of improving enrollment and attendance by decreasing costs of meals and boosting child health by increasing nutrition levels in the served food. The presence of a toilet facility is used to represent the state of school facilities. Toilets are basic necessity amenities that are not always present in schools in developing countries. Finally, the boundary variable measures the presence of an enclosure such as a fence around the school. This serves to represent the safety of the area around a school and the delineation of the educational compound amidst a larger rural or urban area.

Figure 1 shows the variable definitions and basic descriptive statistics on the data used in this paper. Not every block had information on all the variables. Of the 187 blocks, data on completion rates, percentage of trained teachers, and percentage of schools with a lunch program is limited. For the blocks in Karnataka, the mean enrollment rate at 91 percent is higher than the average completion rate at 85.8 percent. It is not surprising that both of these variables are similar in magnitude because of the high correlation between them. With a standard deviation of 9.6 percent, the enrollment variable has a tighter fit than completion.

The descriptive statistics on explanatory variables show increased variability and spread. The average population per village is 1952, but has a large standard deviation. This level of variation will be interesting in this study which seeks to discern any possible size or economies of scale effects associated with educational resource provision for a smaller village compared to a relatively larger village. The other block level characteristic, adult literacy level, stands at 63.5 percent in Karnataka which is slightly higher than the World Bank reported average for India in 2002 at 61.3 percent. Of the explanatory variables, the percentage of trained teachers seems to have the least variability. On average, about 77 percent of the teachers in Karnataka public elementary schools are trained according to NCTE standards. Student teacher ratios which average at about 34 pupils per instructor, with a standard deviation of 17, are slightly under the recommended ratio of 40 students to a teacher which is allocated and funded for under the Sarva Shiksha Amendment to universalize elementary education in India (Sarva Shiksha Abhiyan, 2004). Child population per school, at 1529, seems large because it includes children who are in and out of school. The observations on school inputs and facilities again yield fairly large standard deviations. Thus, an interpretation of these means to be significant in any particular direction is limited by the high variance and spread of the data points.

Figure 1. Descriptive statistics

Variable	Observations	Mean	Standard deviation
Dependent variables			
Enrollment rate	187	0.910	0.096
Completion rate	141	0.858	0.151
Block characteristics			
Population/village	187	1952	1311.6
Adult literacy level / block	187	0.635	0.107
Teacher/classroom characteristics			
% of Trained teachers	178	0.772	0.152
Student teacher ratio	187	33.533	17.027
Child population per school	187	1528.5	678.8
Child pop. per school squared	187	279450	26076
School characteristics			
% of Schools with lunch program	143	0.367	0.288
% of Schools with a toilet facility	186	0.423	0.182
% of Schools with a boundary	187	0.281	0.238

6. Results and Analysis

In this section I will discuss the results of a series of regression specifications. The regressions use either the log odds of enrollment or log odds of completion as the dependent variable. Figure 2 shows the correlations between each of the explanatory variables with the two dependent variables. The first set of estimated models (1) through (8) are shown in Figure 3. These use combinations of the explanatory variables listed under block, teacher / classroom, and school characteristics as regressors. Student-teacher ratios and child populations per school are used as proxies for each other so the variables are included separately in the specifications. The next eight estimated models (9) through (16) are shown in Figure 4 and these incorporate district level fixed effects with dummy variables on the twenty seven districts.

Least squares and range-corrected estimates for the enrollment and completion production functions are presented in Figure 2. The range-corrected results control for the

potential biases associated from estimating regressions that may have predicted values above the truncated range [0,1] for the dependent variables. The log odds, as described in the previous section, of the dependent variables are used in these regressions to minimize the biases in the estimates of variance in the error term. The coefficients listed in the tables are reported as the logged odds ratios. The t-statistic is also provided. A comparison of the regression results (OLS and range-corrected) show that there are no sign changes with the exception of child population per school for enrollment. Additionally, the coefficients that are significant for the OLS regression are also significant for the range-corrected specifications. Due to the biases associated with using standard OLS with a linear probability function, the discussions in this section will be based on the range-corrected results.

Regressions of each of the independent variables on the dependent variable based on the range-corrected estimates show that there are many t-statistics that are significant at the 5 percent level. The block characteristics show mixed significance for completion and enrollment odds. For example, adult literacy level per block seems to have a negligible impact on enrollment, but a strong significant effect for completion odds. The lack of significance of adult literacy on enrollment here does not rule out the possibility of there being an endogenous effect through other explanatory variables. Combining regressors in future specifications will provide insight to this. Population per village seems to have a negative and significant influence on the likelihood that children enroll in school only. This suggests that larger schools may be more likely to discourage enrollment. The same linkage is not apparent for completion rates.

The educational inputs tend to impact enrollment and completion similarly. Notably, percentage of trained teachers per block, with t-statistics of 1.96 for enrollment and -2.75 for completion, and student teacher ratios, with t-statistics of 2.27 for enrollment and -4.04 for

completion, are both found to be significant explanatory variables at the 5 percent level. The child population per school variables attempt to discern economies of scale in the education system, but again the endogeneity issue limits the role these variables have in driving enrollment. A negative coefficient of -2.78 for child population per school on completion odds indicates that having a more crowded school may be a hindrance to a child completing his or her education.

For a number of the school level inputs, the specifications indicate that none of the characteristics significantly impact both enrollment and completion. The percentage of schools with a lunch program is found to be important to enrollment odds with a t-statistic of -2.49. This negative coefficient may serve to indicate that increasing the number of kitchens or presence of lunch programs may actually be counterproductive and detract from the base purpose of education. Parents may take this as a signaling mechanism to not enroll their child in the school. The individual regressions point to the availability of free food being a less prominent factor in driving positive enrollment or completion. Lunch programs are an area where school districts in Karnataka are spending an increasing amount of money and there seems to be little explanatory power between having a free lunch program and enrollment and completion trends. The presence of a boundary, representing safety, and toilet, standing for infrastructure, were both found to be significant for completion odds with t-statistics of -2.57 and 3.43 respectively. The negative coefficient on the toilet facility variable may be a result of high variance in the data or may signify endogenous efforts where the presence of a toilet compounded with large lapses in infrastructure may make the variable significant in a counterintuitive direction.

Figure 2. Individual regressions. Reported coefficient (t-statistic)

Variable	OLS Enrollment	Range-corrected Enrollment	OLS Completion	Range-corrected Completion
Block characteristics				
Population/village	-.625 (-2.11)	-3.793 (-2.12)	-.350 (-.53)	-1.683 (-.67)
Adult literacy level / block	.232 (1.61)	.933 (1.07)	1.553 (6.14)	5.749 (5.91)
Teacher/classroom characteristics				
% of Trained teachers	.412 (3.21)	1.538 (1.96)	-.309 (-1.74)	-2.239 (-2.75)
Student teacher ratio	4.44 (3.65)	17.086 (2.27)	-8.501 (-3.70)	-35.170 (-4.04)
Child population per school	.143 (.3)	-.340 (-.12)	-2.182 (-2.38)	-9.633 (-2.78)
Child pop. per school squared	.001 (1.00)	.003 (.52)	-.001 (-.73)	-.009 (-1.17)
School characteristics				
% of Schools with lunch program	-.072 (.027)	-.376 (-2.49)	-.012 (-.26)	-.057 (-.33)
% of Schools with a toilet facility	-.046 (-.91)	-.115 (-.37)	.337 (3.86)	1.155 (3.43)
% of Schools with a boundary	.055 (2.43)	.436 (3.21)	-.100 (-2.32)	-.402 (-2.57)

Despite the presence of several variables and the size of the sample spanning all the blocks in Karnataka, the undoubted presence of endogeneity in education system minimizes the importance of identifying single policy variables to target in increasing the demand for education. Additionally, the absence of available data at the household level limits the sample size to 187. A glance at the means and standard deviations in the descriptive statistics suggests that there is substantial variation in the data. However, one favorable thing is that there is minimal correlation between the independent variables. A correlation matrix is provided for reference in

Appendix 1. The only notable correlations are between boundary and toilet and student teacher ratio and trained teachers. These will be addressed in later specifications.

Specifications using all regressors except child population per school (1) (2)

OLS estimation of the demand functions with seven regressors are presented in columns (1) and (2) of Figure 3. The first specification (1) has the logged enrollment odds as the dependent variable and includes 149 observations. The estimated coefficients that show significant at the 5 percent level include population per village, student teacher ratio, and percentage of schools with a lunch program and boundary. The effects of the indicators in the aggregate form provide a better explanation for what drives the odds of enrollment. The combination of variables explains about 22.1 percent of the observed variation in enrollment. Re-transforming the logit coefficient on boundary to an odds ratio provides a more meaningful number for discussion. In this specification, a unit increase in the percentage of schools with a boundary leads to a 1.59 unit increase in the odds for enrollment. The magnitude on this variable and its significance with a p-value of .01 suggests that the average perception of safety in a block is pertinent in influencing the likelihood of enrollment. The percentage of schools with a lunch program is also significant with a t-statistic of -2.11. Again, the negative coefficient from the individual measures in Figure 2 carries over to the aggregate form possibly indicating that the provision of free lunch to children may not be as important to the demand for education as other school and classroom inputs. Besides school inputs, there also seem to be higher block level characteristics at work which are embedded in the size of the village. A negative sign on the estimated coefficient for population per village is evidence against a large village being more conducive to higher elementary enrollment.

Figure 3. Regression Results (1) – (8)

Variable	1	2	3	4	5	6	7	8
Constant	2.105	1.973	2.652	2.076	2.515	1.971	1.761	1.184
Block characteristics								
Population/village	-8.105 (-2.94)	-5.685 (-1.94)	-6.611 (-3.50)	1.722 (.79)	-.382 (-1.46)	-6.032 (-2.42)	-9.586 (-2.15)	2.588 (.6)
Adult literacy level / block	1.351 (1.24)	3.628 (3.34)	2.084 (2.23)	3.298 (3.27)	1.571 (1.39)	3.583 (3.35)	1.467 (1.31)	3.301 (3.16)
Teacher/classroom characteristics								
% of Trained teachers	.053 (.06)	-1.730 (-1.62)	.576 (.71)	-1.580 (-1.72)	1.805 (2.26)	-1.965 (-2.52)	1.210 (1.4)	-1.173 (-1.39)
Student teacher ratio	36.393 (.000)	-3.728 (-.32)	26.758 (3.05)	-10.398 (-1.06)				
Child population per school							3.289 (.37)	-25.300 (-2.96)
Child pop. per school squared							.018 (1.38)	.032 (2.67)
School characteristics								
% of Schools with lunch program	-.321 (-2.11)	.013 (.09)			-.358 (-2.26)	.020 (.14)	-.343 (-2.21)	.019 (.13)
% of Schools with a toilet facility	.277 (-.61)	-.142 (-.32)	-.721 (-2.27)	1.006 (3.26)			-3.59 (-.76)	-.011 (-.03)
% of Schools with a boundary	.466 (2.6)	.220 (1.20)	.658 (4.5)	-.349 (-2.43)	.425 (2.61)	.186 (1.21)	.47 (2.44)	.291 (1.62)
Number of observations	149	108	178	132	149	108	149	108
R^2	0.221	0.216	0.197	0.24	0.131	0.2143	0.187	0.286

Regression (2) does a similar logit analysis with seven regressors on the odds of completion. The results are highly disparate compared to specification (1) in that the block level characteristics are the only factors which draw significance. The coefficients for population per village and adult literacy level per block, with t-statistics of -1.94 and 3.34 respectively, draw significance at the 5 percent level. The positive direction of the coefficient on adult literacy points to an increase likelihood for a child to complete school when the average adult in a particular region, in this case a block, is more educated. As children age, the opportunity costs of going to school also increase, but more educated adults appear to be more likely to bypass those pressures and continue to educate their children. Another explanation might just be that the educated adults are the ones earning the higher incomes and therefore find minimal need for their children to work during their school years. The limitation of this data set to a block level analysis prevents further quantitative probing into these areas. The noteworthy school characteristics from specification (1) lose significance when regressed on completion odds. It is difficult to interpret the signs on these coefficients because the 95 percent confidence interval for all seven variables spans a negative to positive range. Due to the fewer available data points on completion rates, this specification only included 108 observations. The smaller sample size may also be another reason why we fail to pick up any significance among the teacher/classroom and school inputs. However, this sample does have a similar R^2 value to specification (1) at 21.6 percent. Perhaps including more data points will help increase the explanatory power of the regressors on completion rate.

Specifications using all regressors except kitchen variable (3) (4)

A larger sample size may provide more explanatory power and fit for the data. Additionally, it may help reinforce (or nullify) some of the observed trends mentioned in previous specifications. Out of the independent variables, the data constriction most lies with the kitchen variable. Specifications (3) and (4) factor in the lack of available data points on the percentage of schools with a kitchen and lunch program for a number of blocks. Since there are only 158 blocks with that observation, this single variable drops the sample size by 19 when included in a regression. Thus, in specification (3) we have the same explanatory variables as specification (1) minus the variable representing lunch programs. This regression now includes 178 observations. By including more sample points, we seek to understand whether there will be a better fit and explanation for the data. For enrollment odds, in regression (3), we see that both variables under block characteristics have now picked up significant at the 5 percent level, in addition to a concentrated effect amongst the coefficients representing schools with toilet facility and boundary. Infrastructure as represented by the toilet facility variable is an important component not only to schools, but also to the perceived characteristics that households take into consideration when making a decision of whether or not to enroll their child in a school. The presence of a boundary appears to be even more significant in this regression than in specification (1) with a t-statistic at 4.5. Interestingly, the R^2 value actually dropped to 19.7 percent when more data points were included. Perhaps a larger sample size actually ends up increasing variation, thereby decreasing explanatory power.

Specification (4) for the odds of completion has an increased sample size from 108 to 132 by dropping the kitchen variable and shows a different set of trends. The notable difference when extracting one variable in specification (4) compared to (2) is that the R^2 value actually

increases to 24 percent, but the odds for completion derives additional explanatory power from the estimated coefficients for the percentage of trained teachers, school with toilet facility, and presence of a boundary. The population per village is no longer a significant variable, perhaps because a bigger sample size captures some of the endogenous effects of block characteristics on classroom and school inputs. For example, a village which is more vocal and prioritizes education may see more apparent changes in school and teacher characteristics. An interpretation of signs on the coefficients is difficult because the confidence interval spans both positive and negative values. The new insight from this specification is that having trained teachers is important to completion, even though we have not observed a similar relationship with this regressor and the odds of enrollment.

Specifications adjusting for high correlations (5) (6)

Specifications (5) and (6) factor the sizeable correlations among a few independent variables. Student teacher ratio and the percentage of teachers who are trained are correlated at a value of .5887 which means that dropping one of these variables might give a better sense of the explanatory power of teacher/classroom inputs on the demand for education. Additionally, the correlation of .5099 between the percentage of schools with a boundary and a toilet also warrants dropping one of the terms. From the individual regressions in Figure 2, we see that the presence of boundaries has slightly higher explanatory power than toilet facilities. Both trained teachers and the student teacher ratios seem to be significant by themselves in driving enrollment and completion. However, the literature suggests that class size is not an important driver of elementary education attendance patterns. Thus, we choose to include the coefficient on training of teachers instead of that for student teacher ratios. Regression (5) drops the student teacher

ratio and toilet variables in the regression with enrollment and does the same for completion in (6). We include the kitchen variable to keep a reasonable number of regressors for analysis.

In specification (5) there were 149 observations and five independent variables regressed on enrollment odds. The t-statistics for trained teachers, kitchen, and boundary measures were 2.26, -2.26, and 2.61 respectively. Again, an interpretation of signs is difficult because the confidence interval spans both positive and negative values. The notable differences when extracting two variables in specification (5) compared to specification (1) is that the percentage of trained teachers is now significant at the 5 percent level. Transforming the logit provides a coefficient interpretation that for each one unit increase in the percentage of trained teachers per school, the odds of enrollment increase by 6.08 units.

Specification (6) for the odds of completion, with a limited set of regressors, shows different patterns. With a t-statistic of -2.52, the coefficient for percentage of teachers who are trained is highly significant with a p-value of .013. The block level characteristics that were significant in specification (2) are still important explanatory inputs in this regression. Overall, specifications (5) and (6) provide a clearer picture of the relationships between the various block, educational, and school inputs and the likelihood for enrollment and completion because they remove the limitations due to multicollinearity. Including both trained and student teacher ratio or boundary and toilet in the same regression equation poses potential downward biases in the t-statistics and upward biases in the standard error.

Specifications using all regressors except student-teacher ratio (7) (8)

In specifications (7) and (8) we use the same base regressors as in regressions (1) and (2) but substitute child population and child population squared as proxies for the student-teacher

ratio. Because Sarva Shiksha Abhiyan funds 1 teacher per 40 students, the child population per school term may better address the endogeneity issue. For example, the policy means that enrollment is supposed to drive the number of teachers, but there is an effect in the other direction as well. Parents may decide to enroll their children in those schools with better teachers resulting in overcrowding. Using child population per school as an instrumental proxy, we try to control for some of this endogeneity. The child population per school squared term is included to capture potential non-linear effects in the explanatory power between the regressors and independent variables. If the relationship is not proportional, forcing a linear relationship will not allow for any non-linear effects and thus might result in further biases in the observed results. The child population per school and the squared term only show significance for completion rates at the 5 percent level. Perhaps the lack of significance for enrollment is an indication that it is not as appropriate of a proxy variable as student teacher ratio to capture the manner in which enrollment is influenced by the number of children and faculty in a school.

The results show that again the odds of enrollment are significantly driven by the availability of a lunch program and boundary and completion is influenced by the adult literacy level per block. The R^2 value of 0.187 for enrollment is slightly lower than specification (1) while the 0.286 value for completion is considerably higher than specification (2). Using the proxy seems to help provide a better fit for completion.

Due to the lack of precision in the results and the potential biases associated with endogeneity when aggregating the regressors for enrollment and completion in the previous regressions, I decided to use a different method for the remaining specifications. Because the observed data points in this study are only available at the block level, rather than at a household

or individual level, it is difficult to assert what drives parental demand without separating out those effects which are driven by community and block level trends. Case and Deaton's (1999) study on school outcomes in South Africa immediately before the end of the apartheid government provides insight into the endogeneity issue associated with community level factors. Under the apartheid regime, resources for Black schools were centrally controlled giving very limited control to the Black population over their location and resource allocation. Therefore, a large fraction of the variation in school resources across districts was independent of the educational choices of Black parents and their children. This is just one example of how the use of an aggregated data set can be limiting because it fails to account for variation between districts. To address this potential bias in the Karnataka block level data set, this section looks at the same eight specifications from (1) to (8) but with district level fixed effects with dummy variables.

Figure 4 shows specifications (9) through (16) with district level dummies. Since decisions and expenditures are made and allocated at the district level in India, this approach is actually extremely insightful. For example, if we look at the variable of percentage of trained teachers we need to consider whether teacher placement is a characteristic of the school or the block. Some districts may care more about their teachers and thus might choose to allocate more resources for their training. The district level fixed approach attempts to explain the variation across samples and hold constant the characteristics of the community. Since this data set does not allow an understanding of the urban-rural differential that might exist in education, a look at possible district clustering effects among the dummy variables may also provide additional insight from the district level fixed effect approach. Endogeneity is an important concern in education and this section looks to minimize the biases associated with it.

Figure 4. Regression Results (9) – (16)

Variable	9	10	11	12	13	14	15	16
Constant	0.87	1.392	1.301	1.459	0.64	1.476	0.062	0.906
Block characteristics								
Population/village	-.412 (-1.2)	1.374 (.4)	-3.353 (-1.5)	1.231 (.49)	-2.309 (-.65)	1.579 (.47)	-3.925 (-.84)	6.780 (1.39)
Adult literacy level / block	-.721 (-.52)	1.217 (1.00)	.731 (.6)	1.212 (1.26)	-.1.065 (-.74)	1.220 (1.02)	-1.230 (-.86)	1.120 (.93)
Teacher/classroom characteristics								
% of Trained teachers	-.736 (1.03)	-.535 (-.5)	-.456 (-.5)	-.612 (-.74)	.225 (.22)	-.429 (-.43)	.267 (.26)	-.341 (-.33)
Student teacher ratio	38.580 (3.28)	3.052 (.25)	27.440 (2.7)	3.492 (.687)				
Child population per school							1.138 (.13)	-16.530 (-1.45)
Child pop. per school squared							.019 (1.38)	.014 (.85)
School characteristics								
% of Schools with lunch program	-.074 (-.29)	-.056 (-.25)			-.093 (-.36)	-.081 (-.37)	-.132 (-.49)	.004 (.02)
% of Schools with a toilet facility	.319 (1.37)	-.278 (-.52)	-.380 (-.87)	-.124 (-.34)			-.601 (-1.01)	-.111 (-.2)
% of Schools with a boundary	.319 (1.37)	.158 (.71)	.289 (1.39)	.133 (-1.37	.222 (1.09)	.098 (.53)	.324 (1.34)	.173 (.79)
Number of observations	149	108	178	132	149	108	149	108
R^2	0.527	0.562	0.531	0.689	0.352	0.56	0.496	0.437

The results from Figure 4 attempt to explain the trends associated with district level fixed effects. Unfortunately, due to the limited sample size, adding on 26 regressors for the district dummies affects the significance of the results. By loosing degrees of freedom, we fail to notice any glaring significant effects in any of the eight regressions except for the estimated coefficient on student teacher ratio on the odds of enrollment in specification (9) and (11). However, the R^2 values have almost doubled and now range from 35.2 percent to 68.9 percent. Removing the variation between districts has given us a better fit for the data.

A look at the clustering effects between districts is also interesting and may be suggestive of some larger forces such as urban-rural differentials at work in this context. The districts which show most significance in their coefficients are 1,5,13,18, and 24 to 26. Appendix 2 includes a district wide map of Karnataka and a table with each district number corresponding to its name. Districts 1 and 5, Bagalkot and Bijapur, are both located in the northern part of the State and both seem to have clustering impact on completion. Clusters for enrollment are most apparent in districts 13 and 18, Bangalore rural and urban, which are found in southeastern Karnataka. Finally, districts 24 to 26, Shimoga, Tumkur, and Udupi have significant clustered dummies for both enrollment and completion. These districts are centrally located. These clumps in the northern, central, and southern portions of Karnataka may be more prone to urban-rural differentials. The literature suggests that the opportunity costs to education are higher for households in rural areas, so these perceived differences in school quality may play a more significant role in explaining the demand for education in those areas. In this study, data is aggregate across urban and rural locations. A more extensive household level study may be more likely to discern these geographic trends and their significance.

7. Conclusions and Policy Implications

In recent years, the expenditures into the educational system in Karnataka have increased to improve access to and quality of schooling. The State has made progress towards universalizing education with unified efforts in addressing concerns for the primary education system between the Government of Karnataka, Sarva Shiksha Abhiyan, and various NGOs. Access to education is no longer as big of a problem in India, but the debate has shifted towards the demand side. The focal point in this study was modeling the average household demand function for schooling. What drives parents to enroll their children in school and keep them there?

Enrollment and completion rates are important determinants of the demand for education. In this study, these two variables were treated as a function of the opportunity costs and benefits of associated with education. These benefits and costs were represented as streams of inputs associated with block, teacher/classroom, and school characteristics. It is difficult to pinpoint single factors in this field because education is a public good and requires a long term investment with a large number and range of factors that influence the decision to attend school. The effect of expected gains or perceived quality through the benefits associated with education is important from a policy perspective because it suggests that investments in educational and school inputs may be used to achieve similar effects as other social reforms aimed at reducing the opportunity costs associated with education. This is particularly true in developing countries, where the average household is much more limited by the availability of resources in making their decisions.

Although limited by data constraints and a small sample size, this study yields some

insightful information on the factors that influence the demand for schooling. Our analysis indicates that block level characteristics play a significant role in effecting the average household's decision to invest in education. For example, size effects associated with population per village hint at some scale economies and the lack thereof to education. For both enrollment and completion, a larger size village was less likely to prompt higher demand for elementary education. Maybe bigger villages in Karnataka tend to prioritize education lower and may need to rethink their overall goals, especially with respect to public goods like education which can easily be sidetracked. Here the government may need to emphasize decentralization policies in villages where there is less emphasis on and ownership of elementary education. The Government of Karnataka's partnership with the Azim Premji Foundation can be leveraged to increase public access to data. This may be instrumental in spearheading a community movement to increase the quality of education.

The adult literacy level per block, in contrast, is a variable that is particularly important to driving the odds of completion. More educated parents are not only more likely to enroll their children in school, but also keep them in school. From a policy standpoint, this suggests that implementing adult literacy programs can have a complementary effect on increasing the number of children who enroll in primary schools and stay until completion.

Under educational inputs, having trained teachers along with classroom and school inputs is important to both enrollment and completion. From a policy standpoint, it appears that investing in teacher training along with other resources is important; otherwise, the system just leads to negative returns from teacher training. The student teacher ratio, as literature suggests, is also less important as a factor than the percentage of teachers that are trained per school.

Having a kitchen for the mid-day meal program seems to have a minimal effect on completion, but more significant role on enrollment odds. The Government of Karnataka has formed a district level implementing committee for the supervision and effective implementation of the Mid-day meal scheme in addition to devoting a significant portion of financial resources. Even though there is a possibility that minimizing the direct costs of attending school by providing free food may improve enrollment and attendance, the marginal dollar expenditure in this area may be more effective in another area of more direct educational inputs.

The other school inputs used for analysis in this study hint at some alternative factors that may go into modeling the household demand function for primary education. For example, the presence of a boundary seems to have a positive effect on enrollment and completion in most cases. Having an enclosure is an indication of safety around a school and also a separation of school property from surrounding areas. The availability of basic amenities such as a toilet facility do not have a consistent effect on enrollment and completion like having a boundary. Overall, the infrastructure as represented three school inputs seems to be more influential on enrollment than completion. Even though the opportunity costs to education increase as children age, perceived benefits such as school inputs seem to be important to getting kids into schools in the first place. Thus, governmental expenditures in these areas are justified and should be continued.

Presently Sarva Shiksha Abhiyan's budgetary allocation for elementary education expenditures in Karnataka seems to mirror the findings from this study. On average, the majority of spending per block, 68 percent, goes to civil works and other school infrastructure needs. The next biggest line item is teacher-related expenditures which total to approximately 20 percent. It is interesting to note that barely a fourth of this spending is for teaching materials.

Both this study and literature in this field emphasize the importance of teaching materials to the classroom and effectiveness of the teacher. The remaining 12 percent is used for administrative and other miscellaneous purposes. With decentralization reforms ramping, this line item can be more focused on creating avenues to increase parental and community involvement in the school. The results from this study can be used by SSA to better allocate expenditures and assign weights to those factors which are most valuable to improving the quality of elementary education n Karnataka.

This study indicates that both perceived benefits and costs are important in modeling the demand function for primary education in Karnataka. The block level analysis gives us unique insight into the factors that are most significant in driving the demand for elementary education at the level where decisions are being made. Benefits seem especially important to enrollment, while completion seems to be effected by both perceived costs and benefits. Decentralization provides the Government with insight into how and why decisions are being made at a local level. This study shows that, since the determinants of enrollment and completion vary, it is especially important for new decentralized reforms in Karnataka to budget expenditures within schools to best address local household concerns in investing in elementary education.

Appendix A. Correlation Matrix

	Enrollment Odds	Completion Odds	Enrollment	Completion	Pop/village	Literacy	% Trained teachers	Student teacher ratio
Enrollment Odds	1							
Completion Odds	0.359	1						
Enrollment	0.8789	0.4079	1					
Completion	0.3974	0.8397	0.5338	1				
Pop/village	-0.1443	-0.1892	-0.1391	-0.2014	1			
Adult literacy	0.1314	0.3019	0.185	0.3075	0.0546	1		
% Trained teachers	0.111	-0.2681	0.2194	-0.1937	-0.0243	-0.038	1	
Student teacher ratio	0.279	-0.2452	0.327	-0.1709	0.3951	0.0051	0.5887	1
Child pop/village	0.0325	-0.2751	0.0427	-0.2504	0.7885	0.001	0.22	0.6902
Child pop/village squared	0.1548	-0.0958	0.1467	-0.003	0.5375	0.038	0.1953	0.5897
Kitchen	-0.242	-0.0631	-0.2449	-0.0839	0.2638	0.2037	0.2158	0.1143
Toilet	0.192	.-073	0.2293	0.1131	-0.035	0.0067	-0.083	-0.0782
Boundary	0.3258	0.0904	0.2622	0.1369	0.1821	-0.0982	-0.2026	0.0445

	Child pop/village squared	Kitchen	Toilet	Boundary
Child pop/village squared	1			
Kitchen	0.158	1		
Toilet	-0.0647	-0.1172	1	
Boundary	0.1591	0.2237	0.5099	1

Appendix B. Districts of Karnataka

District Number	District Name
1	Bagalkot
2	Belgaum
3	Bellary
4	Bidar
5	Bijapur
6	Dharwad
7	Gadag
8	Gulbarga
9	Haveri
10	Koppal
11	Raichur
12	U.K.
13	B. Rur
14	Chamarajnagar

District Number	District Name
15	Kolar
16	Mandya
17	Mysore
18	Bang Urb
19	Chickmagalore
20	Chitradurga
21	D.K.
22	Davanagere
23	Hassan
24	Kodagu
25	Shimoga
26	Tumkur
27	Udupi

Source: www.censusindia.net/.../ karnataka01.html

Bibliography

Azim Premji Foundation: Education MIS. 2005
<http://www.indianngos.com/azimpremjifoundation/emis.htm>

Bedi, Arjun S. and Jeffrey H. Marshall. 1999. "School Attendance and School Enrollment from Rural Honduras." *Economic Development and Cultural Change*. April *47*(3): pp.657-682

Bedi, Arjun S. and Jeffrey H. Marshall. 2002. "Primary School Attendance in Honduras." *Journal of Development Economics. 69*: pp. 129-153

Belfield, Clive R. 2000. *Economic Principles for Education.* Cheltenham, UK: Edward Elgar

Birdsall, Nancy. 1985. "Public Inputs and Child Schooling in Brazil." *Journal of Development Economics*. 18: pp. 67-86

Bliss, James R. 1991. "Strategic and Holistic Images of Effective Schools." in Bliss, James R. William A. Firestone, and Craig E. Richards *Rethinking Effective Schools: Research and Practice*. Englewood Cliffs, New Jersey: Prentice Hall

Carnoy, Martin. 2001. "Are Educational Reforms in Latin America Working? A New Look at Understanding Whether Education is Getting Better." Paper prepared for Inter-American Development Bank, Stanford University

Case, Anne and Angus Deaton. 1999. "School Inputs and Educational Outcomes in South Africa." *The Quarterly Journal of Economics,* August: pp. 1047-1080

Chernichovsky, Dov. 1985. "Socioeconomic and Demographic Aspects of School Enrollment and Attendance in Rural Botswana." *Economic Development and Cultural Change*. January. *33* (2): pp. 319-332

Coleman, James S. 1966. *Equality of Educational Opportunity.* Washington, D.C.: U.S Government Printing Office

Duraisamy, Malathy. 2001. "Demand for and Access to Schooling in Tamil Nadu." in A. Vaidyanathan and P.R. Gopinathan Nair, *Elementary Education in Rural India: A Grassroots View.* New Delhi: Sagar Publications

Friedman, Milton. 1962. "The Role of Government in Education." in *Capitalism and Freedom.* Chapter VI. Chicago and London: University of Chicago Press: pp. 85-107

Glewwe, Paul and Hanan Jacoby. 1994. "Student Achievement and Schooling Choice in

Low-Income Countries: Evidence from Ghana." *The Journal of Human Resources*. Summer *29*(3): pp.843-864

Government of Karnataka. 2003. "Departmental Medium Term Fiscal Plan 2003-04 to 2006-07." *Department of Primary & Secondary Education*. 2003

Government of Karnataka. 2004. Child Census and Household Survey

Hanushek, Eric A. and John E. Jackson. 1977. *Statistical Methods for Social Scientists.* San Francisco: Academic Press

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature.* XXIV. September: pp.1141-77

Hanushek, Eric. 1994. *Making Schools Work: Improving Performance and Controlling Costs.* Washington, D.C: The Brookings Institution

Hanushek, Eric A. 2003. The Economics of Schooling and School Quality: Labor Markets, Distribution and Growth. Vol I. Northampton, MA : Elgar Reference Collection

Hanushek, Eric A. 2003. The Economics of Schooling and School Quality: Efficiency, Competition and Policy. Vol II. Northampton, MA: Elgar Reference Collection

Heyneman, Stephen P. and William A. Loxley. 1983. "The Effect of Primary School Quality on Academic Achievement across Twenty-nine High and Low-Income Countries. *The American Journal of Sociology. 88* (6): pp. 1162 - 94

Hoxby, Caroline M. 2000. :Does Competition Among Public Schools Benefit Students and Taxpayers?" *American Economic Review.* December *90*(5): pp.1209-38

Jain, Sharada. 2004. "In Search of Quality." *Seminar*. April: 27-30

Jamison, T. and Marlaine E. Lockheed. 1987. "Participation in Schooling: Determinants and Learning Outcomes in Nepal." *Economic Development and Cultural Change*. January. *35* (2): pp.297-306

Jhingran, Dhir. 2004. "Beyond Resources." *Seminar*. April: 33-40

Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics*. CXIV. May: pp.497-532

Kuenning, Mary Arends and Sajeda Amin. 2000. "The Effects of Schooling Incentive Programs on Household Resource Allocation in Bangladesh." *Population Council*. No. 133

Lastra, Eduardo F. 2001. “School Effectiveness: A Study of Elementary Public Schools in a Mexican City.” Ph. D. dissertation, Stanford University, March

Manski, C.F. 1989. “Schooling as Experimentation: a Re-Appraisal of the Post-Secondary Dropout Phenomenon.” *Economics of Education Review*. *8*: 305-312

Marshall, JH. 2004. “If You Build it Will They Come? The Effects of School Quality on Primary School Attendance in Rural Guatemala.” Dissertation, Stanford University

Moll, P.G. 1998. “Primary Schooling, Cognitive Skills and Wages in South Africa.” *Economica*. *65*: pp.491 – 505

NCTE Guidelines. 2005. <http://www.ncte-in.org/norms/ele.htm>

Pritchett, Lant and Deon Filmer. 1999. “What Education Production Functions Really Show: A Positive Theory of Education Expenditures.” *Economics of Education Review*. *18*: pp. 223-229

Purkey, S.C. and M.S. Smirth. 1983. “Effective Schools: A Review.” *The Elementary School Journal*. *83*(4): pp. 427-52

Rampal, Anita. 2004. “Unpacking the ‘Quality’ of Schools.” *Seminar*. April: pp.46-52

Sarva Shiksha Abhiyan, 2004 <http://ssa.nic.in/index.asp>

Shariff, Abusaleh and P.K. Ghosh. 2000. “Indian Education Scene and the Public Gap.” *Economic and Political Weekly*. April *35*(16): pp. 1396-1406

Sinha, Amarjeet. 2004. “Is it really possible?” *Seminar*. April: pp. 18-23

Vasavi, A.R. “In the Labyrinth of the Education Bureaucracy.” *Seminar*. April: 30-32

World Bank Indicators.2005. <www.worldbank.org>