

Do IT Service Centers Promote School Enrollment? Evidence from India

Emily Oster*

University of Chicago and NBER

emily.oster@chicagobooth.edu

Bryce Millett Steinberg

Harvard University

mmillett@fas.harvard.edu

May 21, 2013

Abstract

Globalization has changed job opportunities in much of the developing world. In India, outsourcing has created a new class of high-skill jobs which have increased overall returns to schooling. Existing evidence suggests education may broadly respond to this change. We use microdata to evaluate the impact of these jobs on local school enrollment in areas outside of major IT centers. We merge panel data on school enrollment from a comprehensive school-level administrative dataset with detailed data on Information Technology Enabled Services (ITES) center location and founding dates. Using school fixed effects, we find that introducing a new ITES center causes a 5% increase in the number of children enrolled in primary school; this effect is localized to within a few kilometers. We show the effect is driven by English-language schools, consistent with the claim that the impacts are due to changes in returns to schooling, and is not driven by changes in population or income resulting from the ITES center. Supplementary survey evidence suggests that the localization of the effects is driven by limited information diffusion.

Keywords: India, Call Centers, Education, Outsourcing

1 Introduction

In the last thirty years, globalization has dramatically changed job opportunities in the developing world. In many countries this change has increased the skill premium. In India, the focus of this paper, this change has been particularly striking. The number of individuals employed in outsourcing-related businesses has increased from roughly 50,000 in 1991 to over 2 million in 2010 (NASSCOM, 2010); these jobs demand employees with high levels of education and a good command

*Selma Jayachandhran, Robert Jensen, Larry Katz, Andrei Shleifer and participants in seminars at the University of Chicago, Princeton University and Harvard University provided helpful comments. We are grateful to Perwinder Singh for excellent research assistance.

of English, and pay high salaries by Indian standards. The availability of these new opportunities increases the return to education which may, in turn, increase school enrollment, a possibility often floated in popular media (i.e. Giridharadas, 2010).¹ Although India has made significant strides in schooling, even primary school completion is not universal. In the 2005-2006 National Family and Health Survey, for example, only 79.2% of children aged 6 to 14 were enrolled in school.

An academic literature suggests that cities and districts with a major IT presence have experienced changes in education patterns with the growth in these jobs (Munshi and Rosenzweig, 2006; Shastry, 2012). In a randomized evaluation, Jensen (2012) shows that targeted recruitment for these jobs in areas outside major cities also influences schooling choices. Together, these papers suggest that increasing the presence of ITES centers in new areas *could* increase school enrollment in those areas. Without the presence of targeted information campaigns, will the existence of ITES centers in the local area cause parents to change their investment in their children's human capital?

It is this question we address in this paper. To do so, we first estimate the effect of the introduction of these businesses on school enrollment outside of the major IT areas. This allows us to evaluate the validity of the popular claim that these businesses will have broad geographic impacts in all of India. Our data is sufficient to allow us to distinguish the magnitude of impacts over quite small distances, and we argue we are able to make strong causal statements. As we detail below, we find that the impacts of IT centers on school enrollment are large but localized.

Second, with a more qualitative survey we are able to provide some preliminary evidence on the mechanisms behind these effects and their relatively narrow geographic range. We argue this effect is due to limited information dissemination across areas.² This suggests that in the absence of any intervention impacts may *not* be geographically broad, although better information provision about job opportunities could have large impacts.

We begin by using panel data on schooling and the introduction of Information Technology Enabled Services (ITES) firms³ to estimate the impact of these businesses on school enrollment. Our school enrollment data come from a comprehensive administrative dataset and cover three states in India (Karnataka, Andhra Pradesh and Tamil Nadu); each school is observed for a period of four to

¹This question echoes a very large existing literature on the returns to education and school enrollment in both the developing and developed world (e.g. Freeman, 1976; Katz and Murphy, 1992; Heckman, 1993; Kane, 1994; Foster and Rosenzweig, 1996; Griliches, 1997).

²This is consistent with Jensen (2010) and Jensen (2012), both of which suggest interventions which provide better information on job opportunities (in the former case, in a very similar setting) change schooling decisions.

³This is a class of business which includes call centers, as well as data processing, medical imaging and related services.

eight years between 2001 and 2008. We combine this with a newly collected dataset on ITES business locations and founding dates. Our ITES center data includes areas outside of Chennai, Hyderabad and Bangalore, which allows us to estimate the impact of jobs in areas which have not had an overwhelming IT presence. Our ITES center location data allows us to identify the PIN code (similar to a ZIP code) location of each center, which we can link to school location. We use a school fixed effects estimator to analyze how enrollment changes *within an individual school* upon the introduction of a new ITES center to the area.

We estimate the impact of ITES center introduction on schools in the same PIN code and find strong positive effects: the introduction of one additional ITES center to the PIN code is associated with an increase between 4% and 7% in number of children enrolled in the school in the year after the center introduction.⁴ In addition to school fixed effects, we control for several time-varying school infrastructure controls and year fixed effects interacted with state dummies and area demographics. Our preferred specification is one in which we limit to areas most comparable to the areas with ITES centers: areas with English-language schools. This specification yields a coefficient of 4.8%. Our effects are robust to controlling for district specific trends and to limiting to areas which ever have ITES centers.

We next estimate the geographic range of these impacts, and find they are very localized. ITES centers in PIN codes 5 kilometers or less away have a significant positive impact on school enrollment, but it is very small. ITES centers in PIN codes between 5 and 10 kilometers away have no impact.

An issue with interpreting these results causally is the possibility that the introduction of ITES centers anticipates increased school enrollment rather than causing it. The inclusion of school fixed effects in our specification addresses the concern that ITES center introduction is associated with some fixed area characteristic, but they do not address the concern that ITES centers might be introduced to areas which are *changing* more rapidly.⁵ To address this issue, we analyze the impact of ITES centers introduced in future years on current enrollment.⁶ If ITES center operators are targeting areas which have more rapidly increasing schooling, future ITES centers should also correlate with changes in schooling. Similarly, if other variables are changing continuously and

⁴This effect is driven in large part by older children, which is consistent with the fast impact.

⁵We should note that we have no reason to think this type of endogenous placement is common. Conversations with ITES center operators suggested they choose where to locate primarily based on the level of infrastructure and the quality of possible employees, but there was no mention of locating based on anticipated increases in schooling or previous years schooling increases.

⁶This methodology has been used elsewhere to test for similar concerns (Jensen and Oster, 2009; La Ferrara, Chong and Duryea, 2009).

driving both variables, then future ITES centers should correlate with current changes. We do not find evidence for an impact of future ITES centers and the inclusion of the future ITES center measure does not affect our estimate of the impact of current ITES centers.

We also explore whether these impacts vary by language of instruction. ITES jobs almost universally require knowledge of English in addition to high levels of education. Consistent with this, we find that enrollment in English-language schools increases by around 7% with the introduction of each ITES center, whereas there is no significant change for local-language schools.

In terms of magnitude, the results suggest that introducing an ITES center (median size of 80 employees) increases enrollment in the PIN code by 180 children. We can also describe this in terms of share of out-of-school children who enroll. In the three states we use, the nationally representative National Family and Health Survey shows 84.4% of children in this age group reported attending school at all during the 2005-2006 school year. Using our preferred coefficient of 4.8%, this suggests that about 26% of out-of-school children are enrolled as a result of ITES center introduction.

These results point to a causal impact of ITES centers on school enrollment. In robustness checks we address several lingering concerns – that the results reflect changes in population after an ITES center opening, that they reflect changes in income and that they reflect changes in the number of schools – and argue these issues do not drive our results.

In Section 6 we explore mechanisms. We distinguish two possibilities. First, the introduction of an ITES center may impact actual returns to schooling in the local area by providing new jobs at that center; this can explain our results only if labor markets are very localized. Alternatively, it may impact perceived returns to schooling by providing better information about these jobs in general; this can explain our results only if information is very localized.

To distinguish between these possible mechanisms, we conducted a survey in one district in Tamil Nadu (Madurai), which allows us to observe (a) the localization of the labor market and (b) how widely information diffuses. We find evidence in favor of the information story. Our data indicate that many people do travel several kilometers for work, which suggests the narrow geographic range of ITES center impacts is not due to localization of labor markets. In contrast, knowledge is very localized. Even limiting the sample to individuals who live *within one kilometer* of an ITES center we find that those who live closer are more likely to report they know of a center in the local area and to correctly identify what qualifications are required for the job.

In addition to the relationship to the literature on IT development and education discussed

previously, the findings in this paper relate to a large literature on what policies are effective in increasing school enrollment in the developing world (e.g. Duflo, 2001; Kremer, 2003; Kremer et al., 2005; Duflo, Hanna and Ryan, 2012; Burde and Linden, 2009). In this policy space, our results suggest that better information may be effective in promoting school enrollment in areas further from new job options.⁷ We note that although the evidence here focuses on jobs which require high skills, Heath and Mobarak (2012) show evidence that growth in garment factory jobs in Bangladesh also improve schooling for girls, suggesting a similar phenomenon at play even with somewhat lower skill jobs.

The rest of the paper is organized as follows. Section 2 provides some background on ITES centers, and describes the data. Section 3 describes our empirical strategy. Section 4 shows the central results of the paper, and Section 5 discusses robustness. Section 6 presents our survey data and Section 7 concludes.

2 Background and Data

2.1 Background on ITES Centers

Although the concept of “outsourcing” business processes to low-wage countries has been around since the 1970s, the industry remained small until the late 1990s, as time and cost restrictions were large. With the investment in trans-oceanic fiber-optic cables however, the costs of ITES off-shoring plummeted, and with its relatively educated English-speaking low-wage population, India emerged as the dominant provider of business services ranging from call centers to software development.

ITES center jobs are typically high-paying by Indian standards. The average starting salary at such firms is roughly 8,000 Rupees per month (about US\$175), which is almost double the average per capita income of India (Ng and Mitter, 2005). These firms typically come in two types: (1) multinational corporations with subsidiaries or divisions located in India, and (2) Indian “third-party” firms that provide ITES centers and other services for Western companies. Jobs at the Indian firms tend to have lower wages, higher turnover, and less training than the “in-house”

⁷A caveat to this policy argument is that our results hinge on the fact that jobs in ITES centers require additional education; Atkin (2009) finds that growth in the export sector in Mexico actually leads to school *dropout* since export jobs pay well but do not require schooling.

multinational corporation positions (Dossani and Kenney, 2004). The majority of ITES centers are in larger cities such as Bangalore, Delhi, and Mumbai, but they are spreading rapidly to smaller cities all over southern India.

Many of these firms are call centers, which focus on direct telephone interaction with Western customers. Workers make outgoing calls (for services like telemarketing), and take incoming calls (for customer service, tech support, and credit card activation, among other things) for large Western companies. At these centers, “voice” workers conduct calls almost entirely in English, primarily to the United States; thus, fluency is generally a requirement for entry-level positions.⁸ Other “non-voice” business processes outsourced to such firms range greatly in their skill-level, from data entry to software design. English proficiency may be less important for these jobs, but all of the centers in our study report English as a requirement.

From the perspective of this paper, there are at least two central features of ITES centers which we want to highlight. First, they require relatively high rates of education and pay high wages. To the extent that jobs of this type have not been available historically, their existence may well affect the returns to education (both perceived and actual). Second, the vast majority of these jobs require English skills, which is likely to affect the wage returns to learning English.

2.2 Data on School Enrollment

We use a large administrative dataset on primary school enrollment in India called the District Information System for Education (DISE). This dataset has been collected by the Indian government since the late 1990s, although the data used in this paper begins in the early 2000s. Data collection is coordinated at the district level and involves surveys of schools. These school surveys have several parts. First, they collect data on primary school enrollment, including comprehensive data on number of enrolled students by age, grade, gender and caste. These data are designed to reflect enrollment (not attendance) statistics as of September 30th of the school year (which starts in the spring). Second, they collect data on features of the school, including language of instruction and physical plant characteristics. Each school is given a unique ID number, which allows us to follow schools over time.

The area-level survey is less comprehensive and less frequent, but includes some information on village or urban neighborhood characteristics (throughout this paper we will refer to these regions as

⁸Indeed, many of these firms go to great lengths to train their workers to speak with American and British regional dialects, even adopting pseudonyms and memorizing idioms. Some workers report having to watch hours of American television programs to help perfect their speech patterns (Ng and Mitter, 2005).

“neighborhoods”). Most importantly, for most areas in this survey we observe the PIN code location of the school, which allows us to match the area to ITES center locations. A PIN code is similar to a ZIP code in the US; it is smaller than a census block.

The DISE data is collected by the district and then aggregated by each state government. We use data from three states that have been significantly impacted by globalization: Karnataka, Andhra Pradesh and Tamil Nadu.⁹ The number of years of data varies across states. Panel A of Table 1 shows, for each state, the years of data coverage and the number of schools in the first and last year. In later years the dataset is more comprehensive, covering a larger share of schools. Although this means we do not have a balanced panel, by including school fixed effects we ensure that we compare the same schools over time. We use all years of the panel. However, the bulk of our changes are in 2005 and 2006 and we will illustrate our results using these groups. In a robustness section we will show results limited to this group.

Panel B of Table 1 provides some summary statistics on school enrollment and school characteristics. The average school in our sample is fairly small, with 144 students. The physical plant variables indicate schools are not in very good repair. In an average school, only 70% of classrooms are noted to be in good condition by surveyors. Half of the schools report having a boundary wall, half report having electricity and slightly above half have a toilet. Eleven percent of the sample reports at least some instruction in English (this is based on a question about what languages the school teaches in; they could list as many as they wanted).

This data has several limitations. First, as noted, the coverage of our sample differs somewhat across years. In general, the school fixed effects mean this is not a major issue. The one note of caution is that if the schools we observe are different than the schools we do not observe, our results may have limited generalizability. This is unlikely to be a serious issue, however, since our best estimates suggest we cover nearly all schools in India.¹⁰

Second, the DISE data covers only primary schools. It seems plausible, even likely, that much of the impact of ITES centers would be on enrollment in secondary school, since secondary school education is typically necessary for these jobs, and enrollment at that level is lower in general.

⁹These three are also states in which we have a relatively long time series of data. Although there are of course many more areas of India, we argue these areas should be representative of areas most heavily impacted by these jobs.

¹⁰This is actually a somewhat difficult fact to measure. Official statistics on number of schools in India appear to be largely based on the same data we use here so there is no outside source that we can use to verify coverage. The fact that the Indian government uses this as the source of official statistics, however, gives us confidence that we are covering at least an extremely large share of total schools.

Unfortunately, we cannot observe these enrollments; if anything, this may lead us to understate the impacts.

A final important issue is that the data measures total number of children enrolled, not enrollment rates. This leads to the concern that our results reflect changes in population. We discuss this issue in greater detail in Section 5. For a small subset of school years the school also reported the total population of school-aged children in the area. The coverage of these data is limited, but in a robustness check we will use these data and the variable is summarized in Panel B of Table 1.

Although we do not observe enrollment rates, we can use the 2005-2006 National Family and Health Survey to estimate the share of children in this age group who are enrolled; this may be important in exploring what effect size is plausible and which groups are likely to be affected. In the states we consider here, enrollment is highest among children in the 8 to 11 year-old age group (around 92%), and lower for kids aged 6 to 7 (74%) and 12 to 14 (80%). This suggests scope for increases in enrollment among both younger children (due to initial enrollment at younger ages) or among older children (likely due to retention).

2.3 Data on ITES Centers

To match with the data on education, we collected data on ITES centers. We contracted with a firm in India that helps connect Western firms with Indian ITES centers to create a directory of ITES centers in Andhra Pradesh, Karnataka, and Tamil Nadu. They used their contacts, the Internet, and available directories to compile a list of firms, and called each to confirm their existence, the PIN code of their location and their founding date. Our data collection project focused on areas outside of Bangalore, Chennai and Hyderabad, although we did collect some information on centers there. This focus was in line with our desire to estimate the impacts of these firms outside of major IT centers.

This data collection project resulted in a dataset of 401 ITES centers. Figure 1 shows a histogram of ITES center founding dates; the incredible growth in number of centers over time is clear: in our sample, 68% are founded after the year 2000. As we note above, our data on schooling is collected in September for the year spanning June to April, and the ITES center founding dates are given as simply the calendar year of founding. We code the school year 2005-2006 as 2005, and match with ITES centers this way. A school in a PIN code with an ITES center introduced in 2005 is coded as having a new ITES center in the 2005-2006 school year.

The breakdown of number of ITES centers by state is presented in Panel C of Table 1. In

Column 1 we show the count of all ITES centers; Andhra Pradesh is slightly less well-represented, but the number of ITES centers is fairly similar across states. In Column 2 we report these counts for areas outside the major cities of Bangalore, Chennai and Hyderabad (this is the sample we use for analysis). As expected, this limits the sample somewhat and we are left with 260 ITES centers.

Panel D of Table 1 gives a sense of the source of identification we use by showing three categories of schools. Our sample contains roughly 239,000 schools which are in PIN codes which never have ITES centers (or at least not ITES centers we observe). A further 172 schools are in PIN codes which have ITES centers, but do not add ITES centers during the survey period. Finally, we have 408 schools in PIN codes where the number of ITES centers changes over the course of the study. Given that our specifications will include school fixed effects, we are identifying off of these final 408 schools. In some specifications we will limit the comparison group to areas which are more comparable to those which have ITES centers.

In addition to this basic information on ITES center locations and founding dates, we undertook a follow-up survey of the centers in our sample. Although we attempted to survey all centers, in the end we were able to collect data on 83% (the remaining were missed largely due to refusal to answer survey questions). For these centers we have data on whether they operate in English, the number of employees and several employee characteristics. Information on number of employees and English-language operations is available for all the centers we surveyed; demographic information is available for a subset. The ITES centers are relatively small, with a median of 80 employees. Sixty-two percent of the ITES centers handle voice calls in English. Employees are young (median age of 28), largely without children and mostly from the local area.

As a final note, in addition to ITES centers within the same PIN code as the school, we use two variables measuring slightly further centers: those in PIN codes within 5km of the school's own PIN code and those 5-10km away. To calculate distance, we use GPS data on PIN code locations (the latitude and longitude are measured at the post office in each PIN code). We count the number of ITES centers in each of the two neighboring groups.

2.4 Placement of ITES Centers

A central issue in our analysis is the fact that ITES centers are not placed randomly. Our analysis will take advantage of variation over time, so any fixed differences across areas will be adjusted for, but it remains important to understand what drives placement.

We undertake two strategies. First, we can have an initial sense of the magnitude of this threat based on discussion with ITES center operators about location choices. The primary issues they cited when deciding where to locate were infrastructure and transportation: areas with no electricity and roads were not appealing places to operate. In addition, center operators cited their need to find high quality employees cheaply in the local area. There was some sense of a trade-off: there are more qualified individuals in larger cities, but people outside these areas demand lower wages. It is clear that center operators are thinking carefully about cost-benefit considerations. However, the central demographics discussed are very likely to be constant over time, at least over the short time frame of our study.

We are also able to evaluate this endogenous placement statistically using our data. To do this we estimate, at the neighborhood level: (a) the determinants of having an ITES center by the end of the sample in 2007 and (b) the determinants of adding an ITES center during the period we observe. We focus on variables cited by ITES center operators: whether the area has electricity, whether it is in a more urban area and whether there is an English-language school in the area. This last variable is intended to capture the availability of English-speaking individuals. We also include a control for total school enrollment and, in some cases, state fixed effects.

The results from these regressions are shown in Table 2. In general, the results support the interview evidence. More urban areas are more likely to have centers by 2007 and more likely to add them during the sample; these effects hold with and without state fixed effects. Areas with English-language schools are also more likely to have centers and more likely to add them during the sample; again, these results are robust to state fixed effects. We see limited evidence that electricity matters, although this may be due to the high correlation with urbanization; enrollment also does not seem to have any impact.

The inclusion of school fixed effects means that any differences in levels of enrollment associated with these variables will not impact our results. However, if there are differential trends in enrollment across neighborhoods associated with these variables, this could impact our results. To address this, in the results below we will allow for separate year fixed effects for areas that are more urbanized and areas with any English-language schools; this is discussed in more detail below.¹¹

¹¹We do not include separate trends in electricity or initial enrollment level since these do not impact placement; consistent with this lack of impact on placement, including these does not change our results.

3 Empirical Strategy

We estimate the impact of ITES centers on school enrollment using a fixed effects estimator. We observe enrollment in school i in PIN code j at year t ; denote this variable n_{ijt} . In addition, we observe number of ITES centers in PIN code j at year t , which we denote c_{jt} . Our basic regression is shown in Equation (1) below

$$n_{ijt} = \alpha + \beta_1 c_{jt} + \gamma_i + \phi_t + \Psi X_{ijt} + \epsilon_{ijt} \quad (1)$$

where γ_i is a vector of school fixed effects and ϕ_t is a vector of date controls. These date controls include year fixed effects, and year fixed effects interacted with state fixed effects, neighborhood-level electricity, urbanization, school language and whether there is any English-language school in the area. Thus, we allow the year fixed effects to differ by state and by the variables that drive ITES center placement in Table 2. In addition to these fixed effects, we include a set of school-specific time-varying controls (X_{ijt}) measuring school-level infrastructure. The coefficient of interest is β_1 , which captures the effect of ITES centers on school enrollment. This coefficient is identified off of schools in areas which add ITES centers during the sample. Throughout the analysis, we cluster our standard errors at the neighborhood level.¹² We will also estimate this overall regression including district-specific time trends.

Our left-hand side variable is the log of enrollment, allowing us to interpret our results as a percent change in enrollment. The use of the log form leads to the concern that our results could reflect movement between schools of different sizes. For example, if students leave large schools and move to smaller schools the log specification could show an increase even though total enrollment was stable. We address this by weighting our regressions by number of students enrolled in the first year we observe the school, which gives greater weight to larger schools. In addition, in a robustness check we will run these regressions with the level of enrollment (count) on the left hand side.

A concern with estimating this equation on all areas is that our impacts might be identified off

¹²We choose to cluster at the neighborhood level (rather than at the school) since c_{jt} is the same for all schools within a neighborhood-year. In fact, the level of clustering makes relatively little difference – even clustering at the district level gives very similar standard errors. For example: in our primary regression, Column 1 of Table 3, the t-statistic with neighborhood clustering is 2.92 and with district clustering is 3.19, actually larger. In general, the significance of the results we show never change with district clustering. We should note that when we include district-specific trends in the regression we are *not* able to cluster at all given the large number of controls. This means the standard errors are likely biased downward in those regressions, although since the clustering does not make a large difference in general, it seems unlikely this bias is large.

of rural areas which are not at risk of having ITES centers, and these may not be appropriate comparisons for those areas which get ITES centers. Given that, we will estimate, and focus on, a specification in which we limit the sample to areas which have at least one English-language school; these are the areas most “at-risk” for getting ITES centers.¹³ As a further robustness check, we will also limit to areas which ever have ITES centers during the sample period. Although our primary results use fixed effects, in a robustness check we will show our central estimates in first differences.

As noted in the introduction, we are concerned about the possibility that the results are driven by other variables which are changing over time and influence both ITES centers and school enrollment. A related issue is the possibility that ITES center operators consciously introduce centers in places where school enrollment is increasing. To address both of these issues, we estimate whether *future* ITES centers predict current enrollment using Equation (2) below.

$$n_{ijt} = \alpha + \beta_1 c_{jt} + \beta_2 c_{j,t+1} + \gamma_i + \phi_t + \Psi X_{ijt} + \epsilon_{ijt} \quad (2)$$

$c_{j,t+1}$ is a variable measuring number of ITES centers in PIN code j in year $t + 1$. If $\beta_2 > 0$ this would indicate that areas which get ITES centers next year have higher enrollment in this year, relative to their previous enrollment. This would point to ITES centers being introduced in areas which are growing faster. In contrast, a finding that $\beta_2 = 0$ indicates that ITES centers are not introduced into areas in which school enrollment is growing for other reasons. This technique has been used elsewhere to address this concern (Jensen and Oster, 2009; LaFerrara, Chong and Duryea, 2009). We also estimate Equation (2) including a trend for years until a new ITES center is introduced. This allows us to look slightly more generally at whether enrollments are increasing in years up to a new ITES center introduction. It is important to note that this technique does not allow us to rule out the possibility that ITES centers are introduced at exactly the same time as another innovation, and that the other innovation drives school enrollment. However, this possibility seems more remote.¹⁴

One important issue is the coverage of our ITES center dataset. Although we worked to cover as many ITES centers as possible, it seems extremely unlikely that coverage is perfect and there are

¹³This is our strongest predictor of having an ITES center. Virtually all of our ITES centers are located in PIN codes which have at least one English-language school.

¹⁴We also cannot rule out the concern that ITES center operators are targeting areas which seems like they would have large enrollment *responses* to these centers. Under this theory, our results would be valid within sample but would overstate out-of-sample effects. This would require, however, that ITES center operators are choosing locations based on the elasticity of primary school with respect to future returns. This seems unlikely given the difficulty of measuring these parameters and the insignificant effect they would have on short- and medium-run outcomes for the firms. Our conversations with ITES center operators also gave no indication of this type of consideration in placement.

likely areas that have ITES centers that we do not observe. This means that our “control” group of non-changers also contains some schools that should be in the “treatment” group. To the extent that there is a positive effect of ITES centers on school enrollment, this imperfect coverage should bias our estimates of β_1 downward, since the changes in the control group will be more biased upward by the inclusion of “treatment” schools.

4 Results: Impact of ITES Centers on School Enrollment

This section presents our estimates of the impact of ITES centers on enrollment.

4.1 Baseline Results

We begin by illustrating our results for a subset of ITES center introductions in Figure 2. To generate this figure, we focus on four groups of schools: (1) schools that always have an ITES center in their PIN code, (2) schools that add a center between the 2004-05 and 2005-06 school years, (3) schools that add a center between the 2005-06 and 2006-07 school years and (4) schools that never have any ITES centers. Note that this is not the universe of schools which add ITES centers during this period, but it includes many of the changers. We focus on this group because we can isolate a balanced panel of schools which are observed for four years (2004 through 2007). Using this sample of schools, we regress log enrollment on year fixed effects and take the residuals; this removes any consistent year-by-year variation. These residuals are graphed in Figure 2, which shows changes in these residuals relative to the level in 2004.

The key result in Figure 2 is that there are large year-on-year changes in enrollment in the two groups that add ITES centers during the sample, and these changes line up in terms of timing with the ITES center addition. In areas that add a center between 2004 and 2005, schools see a large increase in enrollment between these years, whereas there is only a small increase in schools that always have centers, and no change for schools that add centers later or never add them.¹⁵ Further, for areas that add an ITES center between 2005 and 2006 there is a large increase in enrollment between these years, but no change in the year before. This is the only group with a large increase between 2005 and 2006. Overall, the figure demonstrates large changes in enrollment which

¹⁵As shown in Table 1, most schools never have any ITES centers. For this reason, the year fixed effects are largely identified off of these areas, so when we generate residuals removing these fixed effects, the average residuals in these areas are very close to zero.

correspond to ITES center introductions.¹⁶

In the regressions, we return to our whole panel, which includes areas which add centers in years other than 2005 and 2006. Panel A of Table 3 shows our statistical estimates of the effect of ITES centers on enrollment. Column 1 presents our results using the entire sample. The coefficient on ITES centers is positive and significant: adding one more ITES center increases school enrollment by 4.6%. Column 2 shows this regression with district-specific trends included, to address the concern that districts that have ITES centers introduced are trending differently than those that do not. The coefficient is slightly smaller (3.0%) but still highly significant. Our preferred estimate appears in Column 3, in which we limit to areas with at least one English-language school, which means our non-changer areas are most comparable to the areas which add ITES centers. Although this restricts the sample significantly, the coefficient is similar in magnitude to the overall sample (4.8%) and highly significant. Finally, Column 4 limits further to areas that ever have an ITES center that we observe (including those that change and those that always have a center). The coefficient is again even larger and significant, despite the extreme sample size restriction.

In Panel B of Table 3 we explore whether the introduction of ITES centers in the slightly broader surroundings matter. As described, we do this by estimating the impact of ITES centers in neighboring PIN codes. We focus on those in PIN codes very close by (within 5 kilometers) and those slightly further (5-10 kilometers away).¹⁷ Panel B demonstrates that there are some impacts for ITES centers in the nearest neighbors. Focusing on our preferred specification in Column 3, we find one more ITES center in one of the closest neighboring PIN codes has a significant but very small impact (around 0.2%). ITES centers in more distant PIN codes (5-10 km away) have no impact. This suggests effects are extremely localized.

The evidence in Table 3 suggests a strong connection between ITES centers and total number of children in school. In Online Appendix Table 1 we show these effects broken down by demographic group. We find the effects are similar for girls and boys.¹⁸ We also see similar effects in the three states in the sample, although the effect is significant only in Karnataka and Tamil Nadu. Perhaps most informative, the impacts are largest for older grades. This suggests that much of the

¹⁶It is not clear why those schools which always have a center have an increase and then a decrease in enrollment; this may reflect different yearly conditions in these areas versus those who never have a center.

¹⁷Distance is measured from center-to-center of the PIN codes, so it is possible that a given individual may be closer or further away, but should be in a similar range.

¹⁸The fact that the impact for boys and girls is similar may seem puzzling, given the focus on the female nature of this work. In fact, in the ITES centers in our data, slightly less than half of the employees are women, which may explain the similar impact.

impact may be due to children staying in school rather than newly enrolling at the youngest ages.

In Online Appendix Table 2 we show a number of additional robustness checks. The first panel shows the results in first differences rather than fixed effects. The second shows the effects from Columns 3 and 4 with district trends included. The third panel shows impacts on enrollment levels rather than changes. The fourth panel estimates the impact of having any ITES center, rather than the number of centers.

The results look very similar in all of these robustness checks. An exception is the analysis, in Panel D, of the impact of any ITES center. Although these impacts are positive they are generally not significant. Practically, this is likely due to the fact that we have much less scope for identification in these cases. But, broadly, this suggests that adding more centers matters even if there is already a center around, which may reflect the extreme localization of information – even if there is already a center in your PIN code, getting one closer to you may matter. We will explore this in more detail in Section 6.

In a final analysis, the last panel of this appendix table limits to *only* PIN codes which add a center in 2005 or add a center in 2006, which means the regressions are identified off of the differential timing of changes. In this – essentially our most stringent specification – we still see a large and significant impact of ITES centers on enrollment.

Future ITES Centers

The central threat to the validity of our estimates is the possibility that ITES center introduction anticipates schooling increases rather than causing them. This is related to the issue of endogenous ITES center placement. As discussed above, to the extent that endogenous placement reflects only characteristics which are constant over time this will not drive our results since we include school fixed effects. Further, if trends are different for areas which are urban, or have more English-language schools, we have also addressed this issue. The concern which remains unaddressed in our main specification is the possibility that ITES centers are located in areas that are changing in other ways that we do not observe. There are at least two specific concerns. One is that ITES centers are placed in areas where schooling is increasing more quickly, since center operators are targeting a future labor force (given that our estimates are for primary schools, this would be a fairly distant future). A second concern is that some other unobserved variable (“modernity”, for example) drives both ITES center introduction and school enrollment.

To address this concern directly we estimate whether future ITES center placement predicts current enrollment. If it does, this would suggest ITES center introduction anticipates changes in schooling, rather than causes them. Panel C of Table 3 replicates Panel A, but includes a control for the number of ITES centers in the following year in addition to the indicator for current ITES centers. Adding the control for future ITES centers has only a small impact on our estimates of the effect of current ITES centers. In addition, and more importantly, the effect of future ITES centers is small and not statistically precise, suggesting no strong evidence of pre-trends. In general, we can reject equality between the coefficients on current and future ITES centers.

In Online Appendix Table 3 we do a similar test, but rather than simply controlling for having an ITES center next year, we control for a time trend up to the year of ITES center introduction (the trend is defined so higher values indicate the center introduction is closer in time). If ITES centers are introduced into places where enrollment is increasing more quickly, we should see evidence of a positive trend. We do not see this. The trend coefficients are small and not significant. It is important to note that the results here do not indicate that ITES center placement is exogenous, but instead indicate that this endogenous placement does not drive our results.

4.2 Impacts of ITES Centers by Language of Instruction

The evidence above suggests that overall school enrollment increases in response to ITES center introduction. Here, we turn to separating the result by language of instruction. One of the central features of ITES centers in India is that the vast majority operate in English. In our survey, all of the voice ITES centers (which make up about half of our sample) use English; the majority of non-voice centers also require English. Given this, to the extent that what we observe reflects changes in schooling in response to job opportunities, these changes should disproportionately result in higher English-language school enrollment.

We separate our effects by language of instruction in Table 4. We generate new variables interacting the number of ITES centers with language of instruction and control separately for the impact of ITES centers on local language schools and on English-language schools.¹⁹ Panel A of Table 4 shows our basic test of differences across school types. Column 1 reports impacts on total enrollment using the entire sample. We find the total impact of ITES centers in English-language schools is large and significant; the impact of ITES centers in local-language schools is smaller and

¹⁹The two variables are mutually exclusive; each coefficient can be interpreted as the effect for that school type.

not significant. In our preferred specification (Column 3), we find that enrollment in English-language schools increases by 7.1% for each ITES center introduced. The p-values reported at the bottom of the table indicate we can reject the equality of the impacts in the two school types. One thing which is important to note is that we *do not* see decreases in enrollment in local language schools. The increase in enrollment in English-language schools does not appear to come at the expense of enrollment in local-language schools.

In Panel B of Table 4 we push the data on language further, and separate schools into three groups: those that do not teach in English at all, those that teach some in English and some in another local language and those that teach only in English. Consistent with the larger impact for English-language schools overall, we find the effects are largest for schools that teach exclusively in English. However, the difference between these and those that teach partially in English are small and inconsistently signed. The largest distinction appears to be between schools that teach at least some English and those that teach none.

Finally, in Panel C of Table 4 we test for impacts of future ITES centers separately by language of instruction. Future ITES centers do not impact enrollment in either local-language or English-language schools. The coefficients on future centers are small and inconsistently signed. We should note that we generally cannot reject equality between the coefficients (other than in Column 4); this is due to the more limited precision on the estimates when we separate the schools into the two groups. The p-values do approach marginal significance.

5 Robustness: Changes in Population and Income

This section addresses several key robustness issues. In particular, we evaluate whether it is possible that our results are simply driven by mechanical changes in number of schools, population or income deriving from the ITES center introduction.

5.1 Changes in Population

A key downside of our data on education is that we observe number of students enrolled, not enrollment rates. This introduces the possibility our results could be driven by population increases. The controls thus far address the concern that ITES centers are introduced to more populous areas or areas which are growing faster. However, if the ITES center itself increases population, this could

produce our result. This would be a concern if we were, for example, considering the impact of introducing a large manufacturing plant to an isolated area. In this case, however, we argue this is unlikely to explain more than a very small fraction of the effect we observe.

To begin, it seems appropriate to calibrate the magnitude of our results in terms of the change in number of students. Focusing on our preferred specification in Column 3 of Table 3 we find a 4.8% increase in enrollment after the introduction of an ITES center. Based on a median school size of 143, this is 7 students per school, which aggregates to about 180 students in the PIN code overall.

The first question is whether in-migration among the employees of the ITES center themselves could be driving this change. There are several reasons we think this is unlikely. First, ITES centers tend to employ young, childless individuals. This can be seen in anthropological and ethnographic work on ITES centers in India (i.e. Ng and Mitter, 2005) and directly in our ITES center survey data. In the average center in our sample, managers reported that 10% of employees have children, so the potential increase in children in the area even if all employees were new to the area is small. Second, relocation for work in ITES centers is also relatively rare (12.2% of employees). Even if we assume *all* this relocation is by people with children we find an average of 5.6% employees with children relocate. At the median ITES center, which has 80 employees, this amounts to just 4.4 people with relocated children. This can be compared to the 180 student increase we estimate for the PIN code overall with a single ITES center introduction. In fact, this number is likely to be an upper bound; in reality, the individuals with children are generally the *least* likely to relocate.

There remains a concern that the introduction of an ITES center may bring with it other service jobs, which could increase population.²⁰ This could mean other jobs in the ITES center itself (although this should be captured in our employment measure) or, more likely, jobs working for ITES center employees (e.g. drivers, maids, cooks). If people migrate into the towns for these jobs, this could result in population changes. The first argument against this is again calibration-based. As we note, the total student increase is about 180, and the median ITES center employs 80 people, of whom about 12% migrate in. If we assume that relocation for work in support jobs is as frequent as for work in the ITES centers themselves²¹, then each ITES center employee would have to bring

²⁰It is also possible that the introduction of an ITES center is associated with an overall increase in other types of businesses, which bring in more migrants. The evidence in Panel C of Table 3 above limits this concern; for this to drive our results, it must be the case that these other businesses enter at exactly the same time. This is also largely addressed by our evidence following this paragraph.

²¹In fact, this is an overstatement: based on nationally representative data from India, migration is least likely among low-skill individuals and most likely among those with high skills.

with them, directly or indirectly, 18 children between the ages of 6 and 14. Put differently, if we assume that each ITES center worker hires two new servants, and 10% of those individuals migrate in with two school-aged children, this would explain about 15% of our effect, still very small despite the fact that these are very generous assumptions.

As a second, related, calibration, we note that migration in general in these states in India is fairly limited. In other household survey data (the 1998 and 2005 National Family and Health Survey (NFHS)) we can estimate what share of school-age children report migrating to a new area within the last year. For 84% of clusters in the NFHS sample there are zero in-migrants in the last year among school-aged children.²² Even the clusters at the 90th percentile on this measure still have only 5% of school-aged children who are in-migrants in the last year.

In addition, we show two analyses which test the population mechanism directly. First, in Online Appendix Table 4 we show, for the subset of areas for which the school reports total neighborhood population, the impact of controlling for population on our results (these population data are recorded by the school, and vary across years).²³ We do not want to lean very heavily on the evidence in these regressions since we observe population only for a small subset of the sample and it is unclear how the school estimated population. However, this table demonstrates that including a control for population in the regressions does not significantly impact our estimates. The coefficients are noisier, but this seems to be due to changes in the sample: there is very little difference between Panel A (restricted sample only) and Panel B (restricted sample, with population control).

Second, we note that an important implication of any mechanical mechanism like migration is that larger ITES centers should have larger effects on enrollment. This need not be the case if the effects are driven by changes in information provided by the existence of the ITES center. In fact, we find that the impacts we observe **do not** scale with the size of the ITES center. In Table 5 we show our primary analyses but include in the regression a control for number of employees.²⁴ If enrollment effects were scaling with the size of the ITES center, the coefficient on employees should be positive, but it is negative and indistinguishable from zero in all specifications.

As a final argument, we note that if migration was driving the increase in school enrollment, we

²²In the NFHS a survey cluster typically covers a single village or area within a town and includes a randomly selected subsample of individuals.

²³Since there are multiple schools in each area, we cannot generate enrollment rates off of these data, since the population reported is an area-level population not simply the population relevant for that school. The fact that this is true should also be clear from the coefficient on population; it is much lower than one, which at least partially reflects the fact that as the area population increases, not all of that increase goes to a given school.

²⁴We collected this information for most, but not all, of our ITES centers as detailed in Section 2.

would expect to see similar gains across age groups in schooling, since it would just reflect the ages of the workers' children. As we discussed previously, increases are not homogeneous across age groups, but in fact are larger for older students, which is inconsistent with the population-driven story.

ITES Center Driven Changes in Number of Schools A related, but more minor, concern is that our results are driven by changes in the number of schools in the area. If the introduction of an ITES center causes a decrease in the number of schools then the remaining schools could see enrollment increases even if the total enrollment rate in the area remains constant. We evaluate this by estimating the impact of ITES center introduction on the count of schools in the neighborhood. Estimates are shown in Online Appendix Table 5. The results indicate that changes in school count are not a concern: the impact on number of schools is very small and not significant. We also demonstrate that the introduction of ITES centers does not affect the number of English-language schools or school infrastructure.

5.2 Changes in Income

A second concern with our results is the possibility that ITES centers drive enrollment because they increase income and schooling is a normal good. We note that this seems unlikely, given the results in Table 5, which show no impact of number of employees on enrollment effects, since the total income increase should be greater for larger ITES centers. Still, we take advantage of the fact that existing literature has provided estimates of the income elasticity of school enrollment in similar contexts to estimate the magnitude of predicted enrollment increase resulting from increased income from ITES centers.

A number of papers have estimated the income elasticity of school enrollment in the developing world (Alderman et al., 2001; Glick and Sahn, 2000; Glewwe and Jacoby, 2004; Orazem and King, 2007); these estimates range from 0.25 to 1.25. In Online Appendix A we go through a detailed calculation of the impact of ITES centers on area level income; we estimate an effect of around 0.57%. Combining these figures and comparing to the overall impacts we estimate on school enrollment, the income changes could account for only a small share of the changes in enrollment. Even at the largest elasticity estimate, this figure is only around 15%.

Similar to the case of population, an auxiliary concern is that the ITES center brings other businesses, which also increase income. It is more difficult to rule this out than in the population case. However, the fact that we do not see evidence of pre-trends suggests that these new businesses

would need to arrive at exactly the same time as the ITES centers. In addition, given the very small share of the effect which is plausibly explained by ITES center income, in order for income overall to explain a larger share, these other businesses would need to swamp the ITES centers in their income contribution, which seems unlikely.

It is important to note that our argument in this section is *not* that migration and income changes play no role in our results but simply to say that calibration suggests these effects are small.

6 Mechanisms: Localized Information versus Localized Returns

We draw several conclusions based on the results in Sections 4 and 5. The introduction of an ITES center to an area results in an increase in school enrollment and this increase is concentrated in English-language schools. The observed increase does not appear to be driven by mechanical changes in the number of schools, population or income. Finally, these changes are very localized: ITES centers even slightly further away have little or no impact on enrollment. Based on these results, we argue that the effects we observe reflect responses to changes in the perceived returns to schooling after the introduction of new local job opportunities.

In this section we provide some initial evidence on the mechanisms that drive this effect. We distinguish two possibilities. First, the introduction of an ITES center may impact actual returns to schooling by providing new jobs at that center. Alternatively, it may impact perceived returns to schooling by providing better information about these jobs in general, even if the change in actual job opportunities is limited. This distinction is potentially important for thinking about policy implications. In this section we use a supplementary dataset which we collected in Madurai District (in Tamil Nadu) to provide some evidence on this question.

To fix ideas, consider the simplest model of schooling decision-making in a context with no information frictions. Assume there are two locations, A and B, both of which begin with no ITES centers and otherwise identical job opportunities and education costs. Assume education is a binary choice which carries some positive wage returns. At some date, an ITES center is introduced into area A and (because we are assuming information is shared fully) it is immediately observable to individuals in both areas. The existence of this center increases the wage returns to education while education costs remain the same.

For individuals in area A, the value of education increases by the full amount of the increased

wage returns. For individuals in area B, however, the increase is less because to take advantage of the new jobs, they would need to migrate to area A. Assuming the cost of migration is positive, the reaction of individuals in area B to the ITES center should be smaller than in area A; how much smaller depends on migration costs. Note that these migration costs could be the cost of moving to live in a new area, or the cost of travel to work in that area.

Now consider adding information frictions so the information about the increased returns diffuses only partially (or not at all) between areas A and B. In this case, the response in area B will be less than in area A *even if costs of migration are small*; how much less will depend on how limited information diffusion is. This suggests two pieces of information are key to distinguishing these models: the extent of migration (for work and for children leaving home) and the localization of information.

Survey Data from Madurai

We fielded a survey in Madurai District in Tamil Nadu. Madurai is a small city about 450 kilometers from Chennai with several ITES centers. We surveyed 1000 individuals: 500 in Madurai itself and 250 in each of two smaller towns, Thirumangalam and Peraiyur, 20 to 50 km away. We will focus on the Madurai data for this analysis. We collected data including distance to work, future plans for children and knowledge of ITES centers. Importantly, we collected GPS data on location of households and ITES centers, allowing us to calculate exact distances. Details of the survey appear in Online Appendix B.

Evidence on Costs of Travel and Migration

The Madurai data do not support the claim of very limited migration. The median person in our sample who is working reports working 2 kilometers away from where he lives; 25% work more than 6 kilometers away. Among people with at least ten years of schooling – presumably most likely to work at high wage jobs like those at ITES centers – the median person reports working 4 kilometers away and 25% work more than 10 kilometers away. This suggests that it is not unusual to travel reasonable distances to work.²⁵ Children migrating away from home as adults is even more common. Among children of the sample participants who are over 18, roughly 40% of them live away from

²⁵As a side note, this also supports the argument in Section 5 that our estimates are not driven by population or income. Since people travel for work, any income impacts would be less localized than the ITES center impacts we estimate.

home, and 25% live more than 5 kilometers away.

The evidence on migration is echoed by larger datasets. Data from the National Family and Health Survey show that among working individuals ages 20-35 with at least a secondary school education, roughly 30% have moved in the last five years. Similarly, in the 2001 Census, 29.9% of all persons were living in a town other than that of their birth.

Taking this evidence together, it seems unlikely that our DISE results reflect localized changes in actual returns.

Evidence on Information Diffusion

We turn now to patterns of information diffusion. We focus on relating distance to an ITES center (calculated based on GPS coordinates) to two pieces of data reported by the households: knowledge about ITES centers and whether parents plan on ITES center jobs for their children.

Knowledge of ITES Centers We focus on five knowledge variables: whether the respondent reports knowing anyone who works in an ITES center, whether they report that there is an ITES center within the “local area”²⁶ and three measures of their knowledge about ITES center job qualifications. The job qualification questions listed a set of characteristics (e.g. speak English, college graduate) and asked individuals whether these were “required” for jobs in an ITES center; in some cases the correct answer was yes, and in others it was no. We generate three measures of knowledge: the share of questions for which individuals reported they “didn’t know” whether the qualification was required, the share of the true qualifications they correctly identified and the share of the false qualifications they correctly identified.²⁷ Online Appendix B reports summary statistics on the variables.

We begin by looking at how information varies by distance category. Panel A of Table 6 estimates coefficients on two dummies: being within a half a kilometer of the closest ITES center and being between 0.5 and 1.5 kilometers away; the omitted category is between 1.5 and 3 kilometers away. We see that knowledge is the highest in areas within a half a kilometer of the ITES center on four of the five measures, which is consistent with the evidence from the DISE data of effects decaying over relatively short distances.²⁸ The exception is when we explore impacts on the share of

²⁶The definition of “local” was up to the respondent.

²⁷The correct qualifications were: speak English, use a computer and be a college graduate. The incorrect qualifications were: politically active, have a driver’s license and be a woman.

²⁸Though the negative effect for the middle category (0.5-1.5 km away) is somewhat curious, it is likely being driven by an outlying center. Honeywell, by far the largest ITES center in Madurai, is located on the outskirts of town. Unlike

people who correctly identify true qualifications where nearly everyone gets a perfect score.

Online Appendix Figures 1-5 show smoothed plots of the knowledge outcomes against distance from the closest ITES center. There is strong evidence that information deteriorates quickly in the area right around the ITES center. Between 0 and 2 kilometers, moving further away decreases knowledge. Consistent with the estimate on the second dummy in the regressions in Panel B, there is some evidence that people who live much further away (between 2 and 3 kilometers) have better information.

In Panel B of Table 6 we take this analysis one step further and estimate the impact of distance to an ITES center on knowledge within 1 kilometer. As we squeeze the data in on a smaller area we increase the comparability across individuals, as well as the comparability across the ITES centers to which they are exposed. Despite the smaller sample size, we see a highly significant relationship between distance and knowledge. Individuals who are closer to an ITES center (so distance is smaller) are more likely to report knowing someone who works at one of these businesses, and more likely to report one in the local area. Further, those who are closer to an ITES center are less likely to report they don't know what qualifications are required and more likely to reject the false qualifications.

Child Job Choices Our second piece of evidence on information focuses on job choices for children. In the survey, we asked individuals about the most likely jobs for their child; they were given a list of possible jobs and asked to list three options. We focus on whether they choose the job "Call Center/BPO Worker" as the most likely job and analyze how proximity to an ITES center impacts this outcome. Since enrollment declines as children age, there is more selection in the older sample; given this, we run regressions on the whole sample and limited to children ages 5-10.

Panel C of Table 6 reports regression results from the Madurai-only sample. Column 1 uses the entire sample and estimates coefficients on the two distance dummies; controls are child sex and age, head of household education and whether the respondent reports that "Call Center/BPO worker" is one of the three listed jobs with the highest wages. This first regression shows no evidence that proximity to ITES centers matters for whether parents envision this job for their children. In

the smaller centers, people close to Honeywell all have relatively high basic knowledge of what ITES centers are, and have heard of someone who works there. The deeper knowledge, however, such as qualifications for working there, decays at the same rate as for other ITES centers. This is likely due to the fact that Honeywell is simply much more visible in the neighborhood than a typical ITES center in our sample. Because Honeywell is located relatively far outside the city center, everyone close to it is more than 1.5 km away from any other ITES center. Thus, a large percent of the people who live near enough to Honeywell to be impacted by it actually fall into the "farthest" category in these regressions.

Column 2, however, when we limit to younger children we see a strongly positive impact of being close to an ITES center.

The difference across age groups could reflect differential selection. It is also possible that this difference reflects the fact that schooling choices are more malleable for younger children – for example, it might still be possible to switch them to an English school. To explore this, in Columns 3 and 4 we interact distance with whether the child is enrolled in an English-language school (controlling for the overall English-language impact). These results are more striking. For both the overall sample and for the younger children, we observe that for children enrolled in an English-language school, proximity to an ITES center strongly impacts whether the parent reports that an ITES center job is likely. The fact that this occurs for the overall sample in addition to the younger children suggests that the lack of impact for the total sample in Column 1 is due to lack of language flexibility among older children.

Overall, this table suggests that there is an increase in perceived chance of ITES center jobs for children when an ITES center is closer. Again, this points to a very sharp decay of information about these jobs even over small distances.

Returns to Schooling and Reported Changes in Behavior As a final note, we present two more speculative pieces of evidence that are supportive of the information story. The first is on returns to schooling. We asked individuals their “best guess” about the monthly wage in the area for someone with a secondary school degree and for someone with only primary school; we define “returns to schooling” as the simple difference between these two values. Areas within half a kilometer of an ITES center report monthly returns 350 Rs higher than more distant areas and this effect is significant (tables available from the authors).

The second piece of evidence comes from the last question on the survey. For the 131 individuals in the sample who reported knowing of an ITES center in the local area, we asked whether they had made any change in response to that center introduction. Of course, it is extremely difficult to interpret responses to questions like this, especially given that it was asked at the end of the survey, which leaves open concerns about priming. However, the results are striking. About 50% report intentions to increase schooling for their children, some of whom cite specifically that they will enroll their children in English-language schools. It is interesting to note that this is the only behavior change reported – there is no mention of individuals themselves getting jobs at ITES centers – which is consistent with the evidence in Section 5 that these centers probably do not

have large impacts on current income.²⁹

We argue that the evidence in this section suggests that the localized impact of ITES centers that we observe in the DISE data reflect limited information diffusion rather than localized labor markets. It is worth noting that, in addition to being complementary to our findings above, this evidence is also quite complementary with experimental evidence on the role of information in schooling – in particular, Jensen (2010) and Jensen (2012). Both of those papers suggest that experimentally varying information about returns to schooling impacts schooling choices, and that is true even though in principle the information is available. The evidence here indicates that very limited information diffusion in the non-intervention context may explain why a simple provision of information can be so powerful.

7 Conclusion

In this paper we argue that the introduction of ITES centers in India has large impacts on school enrollment, and these impacts are concentrated in the very local areas around the ITES centers. We argue this effect is causal, and is not driven by pre-trends or mechanical changes in area-level population or income. The very local nature of our analysis and the fine timing of the effects are helpful in ruling out the concern that endogenous placement or trends in unobservables drive the impacts we see. Further, we provide some suggestive evidence that the very localized nature of the impacts may reflect limited information about non-local job opportunities; we argue this is more likely than the claim that these new job opportunities only impact local returns to schooling.

It may seem puzzling on its face that information is so localized given that people do travel for these jobs. Although it is beyond the scope of our data to prove this, it seems plausible that the type of people who actually have these jobs, and are therefore traveling for them, are not on the margin with respect to their childrens' school enrollment. It is likely to be parents and families who are less well off who are on this margin, and these may be individuals with more limited travel options and more limited exposure to others who work in these sectors. It seems plausible that, in fact, it is only

²⁹Of 131 individuals who know of an ITES center in their local area, when asked if they will change their behavior because of the ITES center, 47 answered no, 9 answered that they will make their child study more or longer, 19 answered that they will make their child learn English, 38 answered that they will make their child learn computer skills or typing, and 18 indicated that there will be a change, but do not specify further. Not one parent answered that they would try to get a job at this center, or a nearby business, or any other change that was not related to investment in human capital for their child. While there is some danger of priming with this question (it was asked at the end of a survey about education and ITES centers, among other things), it is consistent with our assertion that these ITES centers are causing increased enrollment directly, through information dispersion about returns to schooling.

having a physical center in a nearby location that provides information to this marginal group.

It is worth discussing the magnitude of our results, both in general and compared to other interventions to promote schooling in the developing world. Our preferred coefficient indicates that an additional ITES center prompts a 4.8% increase in school enrollment. Based on the National Family and Health Survey, a nationally representative survey run in 2005-2006, 84.4% of children aged 6-14 in our states attended school at any time in the previous year. Our coefficient implies that 25.6% of out-of-school children would be prompted to enroll by an ITES center.

Put differently, our estimates imply about a 4.1 percentage point increase in the enrollment rate. This number is comparable to enrollment effects of other interventions designed to increase schooling in the developing world. For example, the conditional cash transfers in PROGRESA increased schooling 3.4-3.6 percentage points (Schultz, 2004). A program in Kenya which provided school uniforms to girls in Kenya (worth about 1.75% of average yearly income) increased enrollment by 6 percentage points (Evans, Kremer and Ngatia, 2008). Miguel and Kremer (2004) found that administering deworming drugs decreased absence by 7 percentage points, although they do not report effects on enrollment. Our coefficient is similar to (although slightly smaller than) the 5.2 percentage point increase that Jensen (2012) identifies as a response to call center recruitment services for women.

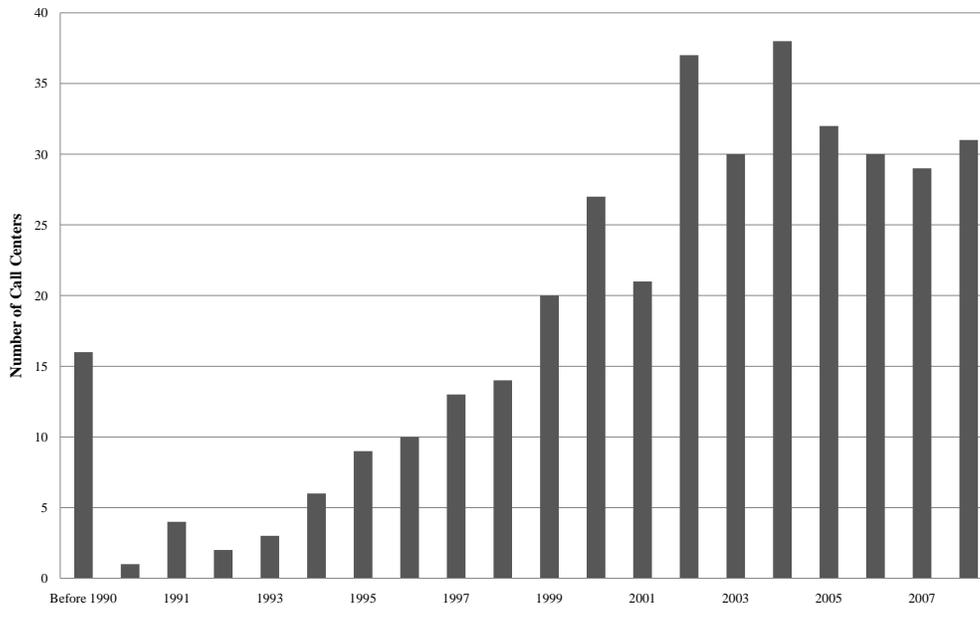
From a policy standpoint, the results provide support for interventions which inform students about returns to schooling (as in Jensen, 2010 and 2012). In the absence of this type of policy, we would expect short-term gains in enrollment to be concentrated around areas with local ITES centers; the evidence in Section 6 suggests this concentration could be limited by broader information sharing.

References

- Alderman, Harold, Peter Orazem, and Elizabeth Paterno**, “School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan,” *Journal of Human Resources*, 2001, 36 (2), 304–326.
- Atkin, David**, “Endogenous Skill Acquisition and Export Manufacturing in Mexico,” *mimeo*, Yale University, 2009.
- Burde, Dana and Leigh L. Linden**, “The Effect of Proximate Schools: A Randomized Controlled Trial in Afghanistan,” *Mimeo*, Columbia University, December 2009.
- Dossani, Rafiq and Martin Kenney**, “The Next Wave of Globalization? Exploring the Relocation of Service Provision to India,” *Working Paper 156*, 2004.
- Dufo, Esther**, “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment,” *American Economic Review*, 2001, 91 (4).
- , **Rema Hanna, and Stephen Ryan**, “Incentives Work: Getting Teachers to Come to School,” *American Economic Review*, 2012, 102 (4), 1241–1278.
- Evans, David, Michael Kremer, and Muthoni Ngatia**, “The Impact of Distributing School Uniforms on Children’s Education in Kenya,” *mimeo*, World Bank, 2008.
- Foster, Andrew and Mark Rosenzweig**, “Technical Change and Human Capital Returns and Investments: Evidence from the Green Revolution,” *American Economic Review*, 1996, 86 (4), 931–953.
- Freeman, Richard**, *The Overeducated American*, New York: Academic Press, 1976.
- Giridharadas, Anand**, “The Caste Buster,” *New York Times*, December 30, 2010.
- Glewwe, Paul and Hanan G. Jacoby**, “Economic growth and the demand for education: is there a wealth effect?,” *Journal of Development Economics*, 2004, 74 (1), 33–51.
- Glick, Peter and David E. Sahn**, “Schooling of girls and boys in a West African country: the effects of parental education, income, and household structure,” *Economics of Education Review*, 2000, 19, 63–87.
- Griliches, Zvi**, “Education, Human Capital, and Growth: A Personal Perspective,” *Journal of Labor Economics*, 1997, 15, S330–S344.
- Heath, Rachel and Mushfiq Mobarak**, “Does Demand or Supply Constrain Investments in Education? Evidence from Garment Sector Jobs in Bangladesh,” *Mimeo*, Yale University, 2012.
- Heckman, James**, *Private Sector Skill Formation: International Comparisons*, University of Chicago Press, 1993.
- Jensen, Robert**, “The (Perceived) Returns to Education and the Demand for Schooling,” *Quarterly Journal of Economics*, 2010, 125, 515–548.
- , “Do Labor Market Opportunities Affect Young Women’s Work and Family Decisions? Experimental Evidence from India,” *Quarterly Journal of Economics*, 2012, 127 (2), 753–792.

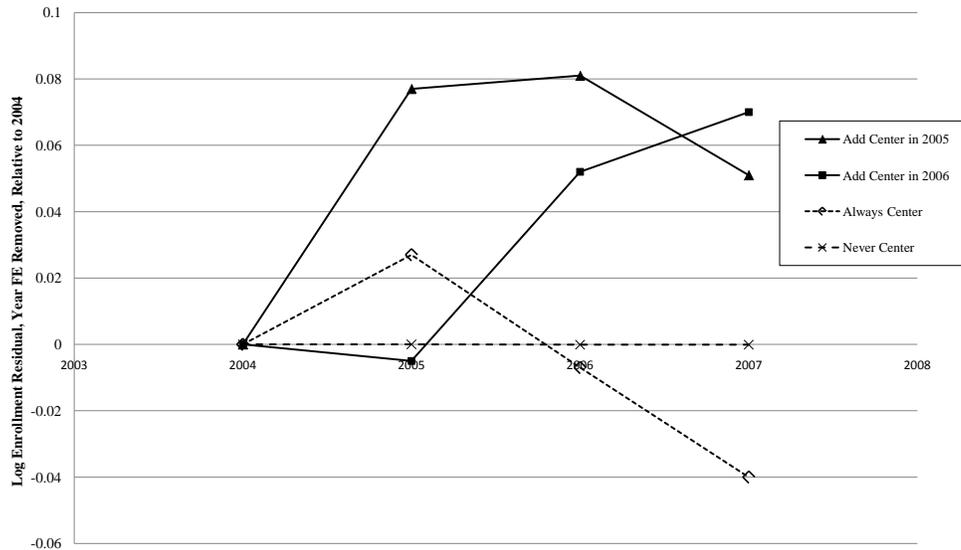
- **and Emily Oster**, “The Power of TV: Cable Television and Women’s Status in India,” *Quarterly Journal of Economics*, 2009, *124* (3), 465–476.
- Kane, Thomas**, “College Entry by Blacks since 1970: The Role of College Costs, Family Background and the Returns to Education,” *Journal of Political Economy*, 1994, *102* (5), 878–911.
- Katz, Lawrence and Kevin Murphy**, “Changes in Relative Wages, 1963-1987: Supply and Demand Factors,” *Quarterly Journal of Economics*, 1992, *107* (1), 35–78.
- Kremer, Michael**, “Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons,” *American Economic Review*, 2003, *92* (2).
- , **Nazmul Chaudhury, Halsey Rogers, Karthik Muralidharan, and Jeffrey Hammer**, “Teacher Absence in India: A Snapshot,” *Journal of the European Economic Association*, 2005, *3* (2-3).
- LaFerrara, Eliana, Alberto Chong, and Suzanne Duryea**, “Soap Operas and Fertility: Evidence from Brazil,” *American Economic Journal: Applied Economics*, 2009.
- Miguel, Edward and Michael Kremer**, “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities,” *Econometrica*, 2004, *72* (1), 159–217.
- Munshi, Kaivan and Mark Rosenzweig**, “Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy,” *American Economic Review*, 2006, *96* (4).
- NASSCOM**, “Strategic Review 2010,” *New Delhi: National Association of Software and Service Companies*, 2010.
- Ng, Cecelia and Swasti Mitter**, “Valuing Women’s Voices: Call Center Workers in Malaysia and India,” *Gender and Technology Development*, 2005, *9*, 209–233.
- Orazem, Peter and Elizabeth M. King**, “Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy,” *Iowa State University, Department of Economics Working Paper 07019*, 2007.
- Schultz, Paul**, “School Subsidies for the Poor: Evaluating the Mexican PROGRESA Poverty Program,” *Journal of Development Economics*, 2004, *74* (1), 199–250.
- Shastry, Gauri Kartini**, “Human Capital Response to Globalization: Education and Information Technology in India,” *Journal of Human Resources*, 2012, *47* (2), 287–330.

**Figure 1:
Distribution of ITES Center Founding Dates**



Notes: This figure shows the distribution of ITES center founding dates among centers in our sample.

**Figure 2:
Impact of ITES Centers on School Enrollment**



Notes: This figure shows changes in enrollment over time for four balanced panels of schools. All enrollment numbers are residuals from a regression of log enrollment on year fixed effects and values represent changes relative to the residual values in 2004. The year 2004 refers to the 2004-2005 school year (beginning in June 2004, ending in April 2005); enrollment is recorded as of September 2004. Schools are coded as adding a center in 2005 if an ITES center is founded in the area any time during 2005.

Table 1: **Summary Statistics**

Panel A: Years of Coverage and Number of Schools			
	Andhra Pradesh	Karnataka	Tamil Nadu
Years of Data Coverage	2004-2007	2001-2007	2003-2007
Number of Schools in:			
First Year of Coverage	59,121	27,136	43,662
Last Year of Coverage (2007)	98,485	52,369	51,548
Panel B: School Summary Statistics			
	Mean	Std. Dev.	Observations
Total Enrollment	143.8	166.4	905,838
% Classrooms in Good Condition	70.7	37.2	905,838
% Schools with Electricity	49.0	50.0	905,838
% Schools with Boundary Walls	51.3	50.0	905,838
Teach in English (0/1)	0.11	0.32	905,717
Total School-Age Population	163.1	1,401	255,355
Panel C: Number of ITES Centers By State			
	Number of ITES Centers		
	Including Cities	Excluding Cities	
Andhra Pradesh	100	74	
Karnataka	144	121	
Tamil Nadu	157	65	
Panel D: Number of Schools by Category			
	Number of Schools		
Never Had an ITES Center	238,986		
Has Same Number of ITES Centers	172		
Has Change in Number of ITES Centers	408		

Notes: This table shows summary statistics for our sample of schools and ITES centers. Panel A shows years of data coverage and summary statistics by state for the three states in our data set. Panel B shows summary statistics on enrollment and school characteristics for the sample of schools used in the analysis. Population is recorded by the schools for only a subset of schools and years. Panel C shows the number of ITES centers in our sample for each state, including and excluding cities (all later analysis will exclude the major cities of Hyderabad, Chennai, and Bangalore). Location (PIN code) and founding year were collected in a primary survey; only centers with both location and founding date were included in the sample. Panel D shows the number of schools in our analysis which are matched to PIN codes which ever have an ITES center in our data.

Table 2: Placement of ITES Centers

<i>Dependent Variable:</i>	<i>Number of ITES Centers, 2007</i>		<i>Add ITES Center During Sample</i>	
	(1)	(2)	(3)	(4)
Ever Had Electricity	-.0002 (.0005)	-.0001 (.0005)	-.0002 (.0003)	.0001 (.0003)
Urban	.010*** (.001)	.010*** (.001)	.004*** (.001)	.004*** (.001)
Any English School (0/1)	.007*** (.0008)	.006*** (.0008)	.003*** (.0004)	.003*** (.0004)
Log Enroll. First Survey Year	-.00005 (.0001)	-.0001 (.0002)	.000001 (.0001)	-.0001 (.0001)
State Fixed Effects	NO	YES	NO	YES
R-squared	0.002	0.003	.002	.002
Observations	71,667	71,667	71,890	71,890

Notes: This table shows the effects of PIN code characteristics on ITES center placement. The left hand side variable in Columns 1 and 2 is the number of ITES centers in 2007; in Columns 3 and 4 it is whether any centers were added during the sample period. Standard errors in parentheses, clustered at the PIN code level. *significant at 10% **significant at 5% ***significant at 1%.

Table 3: Effect of ITES Centers on School Enrollment

<i>Dependent Variable:</i>			<i>Log Enrollment</i>	
Panel A: Number of ITES Centers in PIN Code				
Sample:	All Schools		In PIN Code with Any English Schools	Ever Had an ITES Center
<i>Controls:</i>	<i>Standard</i>	<i>District Trends</i>	<i>Standard</i>	<i>Standard</i>
	(1)	(2)	(3)	(4)
Number of ITES Centers	.046*** (.016)	.030*** (.010)	.048*** (.015)	.071*** (.014)
Observations	911,499	911,499	314,476	2,121
Panel B: Number of ITES Centers in Neighboring PIN Codes				
Sample:	All Schools		In PIN Code with Any English Schools	Ever Had an ITES Center
<i>Controls:</i>	<i>Standard</i>	<i>District Trends</i>	<i>Standard</i>	<i>Standard</i>
	(1)	(2)	(3)	(4)
ITES Centers in PIN Code	.038*** (.011)	.024** (.010)	.040*** (.012)	.075*** (.018)
# ITES Centers in PIN Codes within 5km	.002** (.0007)	.002*** (.0005)	.002** (.001)	.002** (.0006)
# ITES Centers in PIN Codes 5-10km away	.0002 (.0004)	-.0002 (.0002)	.0002 (.0004)	.00003 (.001)
Observations	911,499	911,499	314,476	2,963
Panel C: Impact of Future ITES Centers				
Sample:	All Schools		In PIN Code with Any English Schools	Ever Had an ITES Center
<i>Controls:</i>	<i>Standard</i>	<i>District Trends</i>	<i>Standard</i>	<i>Standard</i>
	(1)	(2)	(3)	(4)
ITES Centers	.046*** (.016)	.033** (.014)	.050*** (.017)	.072*** (.019)
ITES Centers Next Year	-.001 (.013)	-.004 (.012)	-.003 (.014)	-.001 (.016)
p-value, This Year=Next	.06	.11	.05	.02
Observations	911,499	911,499	314,476	2,121

Standard controls: School fixed effects, time-varying school plant characteristics, year dummies interacted with dummies for state, urban, school language, and English language school in PIN code.
District Trend controls: Standard controls plus district-specific trends.

Notes: This table shows our primary estimates of the impact of ITES centers on school enrollment. The independent variable measures the number of ITES centers in the same PIN code as the school. Columns 1-2 include all schools. Column 3 is limited to PIN codes with any English schools. Column 4 is limited to schools which ever have an ITES center in their PIN code (either always have the same number or change during the sample). Standard errors (in parentheses) are clustered at the neighborhood level in Columns 1, 3 and 4; clustered errors could not be estimated when district trends are included in Column 2. *significant at 10% **significant at 5% ***significant at 1%. All regressions are weighted by initial school enrollment level.

Table 4: **Effects by Language of Instruction**

<i>Dependent Variable:</i>		<i>Log Enrollment</i>		
Sample:	All Schools	In PIN Code with Any English Schools	Ever Had an ITES Center	
Panel A: Impact of ITES Centers by School Language of Instruction				
<i>Controls:</i>	<i>Standard</i>	<i>District Trends</i>	<i>Standard</i>	<i>Standard</i>
# Centers × Local Lang.	.019 (.019)	.004 (.015)	.020 (.021)	.044** (.018)
# Centers × English	.073*** (.016)	.057*** (.015)	.071*** (.017)	.102*** (.020)
p-value, English=Local Language	0.01	0.01	0.03	0.02
Observations	911,499	911,499	314,476	2,121
Panel B: Impact of ITES Centers by Detailed School Language of Instruction				
<i>Controls:</i>	<i>Standard</i>	<i>District Trends</i>	<i>Standard</i>	<i>Standard</i>
# Centers × Local Lang.	.019 (.019)	.004 (.015)	.020 (.021)	.045** (.018)
# Centers × Some English	.072*** (.0169)	.061*** (.015)	.069*** (.018)	.096*** (.024)
# Centers × All English	.074*** (.030)	.051** (.023)	.073*** (.032)	.118*** (.033)
Observations	911,499	911,499	314,476	2,121
Panel C: Pretrends by Language				
<i>Controls:</i>	<i>Standard</i>	<i>District Trends</i>	<i>Standard</i>	<i>Standard</i>
# Centers × Local Lang.	.031 (.021)	.018 (.019)	.036 (.024)	.053** (.019)
# Centers × English	.065*** (.023)	.052*** (.020)	.065*** (.023)	.097*** (.032)
# Centers Next Year × Local Lang.	-.018 (.017)	-.019 (.018)	-.022 (.020)	-.017 (.016)
# Centers Next Year × English.	.009 (.023)	.005 (.016)	.007 (.023)	.007 (.022)
p-value this Year vs. Next, Local	0.11	0.26	0.11	0.02
p-value this Year vs. Next, English	0.20	0.16	0.18	0.09
Observations	911,499	911,499	314,476	2,121
<i>Standard controls: School fixed effects, time-varying school plant characteristics, year dummies interacted with dummies for state, urban, school language, and English language school in PIN code.</i>				
<i>District Trend Controls: Standard controls plus district-specific trends.</i>				

Notes: This table shows the impact of ITES centers by school language of instruction. Panel A shows the differential effects for English and local language schools. Panel B shows the effect for local language schools, schools with some English instruction, and schools with exclusive English instruction. Panel C shows the effects by school language for voice and non-voice ITES centers. Voice centers are defined as ITES centers where at least half of employees handle voice calls. Columns 1-2 include all schools. Column 3 is limited to PIN codes with any English schools. Column 4 is limited to schools which ever have an ITES center in their PIN code. Standard errors (in parentheses) are clustered at the PIN code level in Columns 1, 3 and 4; clustered errors could not be estimated when district trends are included in Column 2. *significant at 10% **significant at 5% ***significant at 1%. All regressions are weighted by initial school enrollment level.

Table 5: **Robustness: Number of Employees**

<i>Dependent Variable:</i>			<i>Log Enrollment</i>	
<i>Sample:</i>	All Schools		In PIN Code with Any English Schools	Ever Had an ITES Center
<i>Controls:</i>	<i>Standard</i>	<i>District Trends</i>	<i>Standard</i>	<i>Standard</i>
	(1)	(2)	(3)	(4)
Number of ITES Centers	.073*** (.017)	.062*** (.016)	.069*** (.017)	.099*** (.018)
Log Number of Employees	-.0002 (.0001)	-.0002*** (.0001)	-.0001 (.0001)	-.0002** (.0001)
Observations	911,499	911,499	314,476	2,121

Standard controls: School fixed effects, time-varying school plant characteristics, year

dummies interacted with dummies for state, urban, school language, and English language school in PIN code.

District Trend Controls: Standard controls plus district-specific trends.

Notes: This table shows the effect of the number of employees of a call center on school enrollment. The independent variables measure the number of ITES centers in the same PIN code as the school, and the natural log of the number of employees in the center, respectively. Columns 1-2 include all schools. Column 3 is limited to PIN codes with any English schools. Column 4 is limited to schools which ever have an ITES center in their PIN code (either always have the same number or change during the sample). Standard errors (in parentheses) are clustered at the PIN code level in Columns 1, 3 and 4; clustered errors could not be estimated when district trends are included in Column 2. *significant at 10% **significant at 5% *** significant at 1%.

Table 6: Knowledge of ITES

Panel A: Knowledge of ITES by Distance to Closest, Within Madurai					
	Heard of Someone who works at ITES	Know of ITES in local area	% ITES Ques. “Don’t Know”	% True Qual. Answer Correct	% False Qual. Answer Correct
ITES <0.5 km	.212*** (.082)	.230*** (.078)	-.099** (.042)	-.015 (.042)	.265*** (.061)
ITES 0.5-1.5 km	-.129*** (.040)	-.110*** (.037)	-.009 (.041)	-.021 (.021)	.015 (.030)
Controls	YES	YES	YES	YES	YES
Observations	494	496	498	498	498
Panel B: Knowledge of ITES by Distance to Closest, Within 1 km of an ITES Center					
	Heard of Someone	Know of ITES	% “Don’t Know”	% True Correct	% False Correct
Distance to ITES	-.641*** (.130)	-.488*** (.131)	.215*** (.065)	-.052 (.070)	-.445*** (.111)
Controls	YES	YES	YES	YES	YES
Observations	137	139	139	139	139
Panel C: ITES as Job Possibility for Child					
<i>Dependent Variable: ITES listed first as possible job for child</i>					
	All Children	Children Ages <10	All Children	Children Ages <10	
<0.5 km to ITES	-.004 (.065)	.167** (.080)	-.072 (.075)	-.035 (.095)	
<1 km to ITES	-.0008 (.039)	.026 (.042)	-.063 (.049)	-.024 (.051)	
In English School			-.100* (.057)	-.092 (.060)	
<0.5 km ×Eng.			.292* (.153)	.586*** (.168)	
<1 km ×Eng.			.175** (.081)	.129 (.082)	
Observations	170	81	169	81	
<i>Controls: Child age, child sex, head of household education, whether respondent reports call center as highest wage job, asset ownership (television, radio, refrigerator, toilet).</i>					

Notes: Data comes from the survey run in Madurai District, Tamil Nadu. All regressions are limited to households within Madurai. In Panels A and C, the omitted distance category is more than 1.5 kilometers away. Dependent variables are the same in Panels A and B (abbreviated in Panel B). Columns 3-5 rely on answers to a set of six questions about qualifications which are “required” for job in an ITES center. Details are in Section 6. Controls: head of household education, whether the household head speaks any English and number of assets held by household (television, radio, refrigerator, and toilet). Standard errors in parentheses. *significant at 10% **significant at 5% ***significant at 1%.