WPS5674

Policy Research Working Paper 5674

Students Today, Teachers Tomorrow?

Identifying Constraints on the Provision of Education

Tahir Andrabi
Jishnu Das
Asim Ijaz Khwaja

The World Bank
Development Research Group
Human Development and Public Services Team
June 2011

Policy Research Working Paper 5674

Abstract

With an estimated 115 million children not attending primary school in the developing world, increasing access to education is critical. Resource constraints limit the effectiveness of demand-based subsidies. This paper focuses on the importance of a supply-side factor—the availability of low-cost teachers—and the resulting ability of the market to offer affordable education. The authors first show that private schools are three times more likely to emerge in villages with government girls' secondary schools (GSS). Identification is obtained by using official school construction guidelines as an instrument for the presence of GSS. In contrast, there is little or no relationship between the presence of a private school and girls' primary or boys' primary and secondary government schools. In support of a supply-channel, the authors then show that, for villages that received a GSS, there are over twice as many educated women and that private school teachers' wages are 27 percent lower in these villages. In an environment with poor female education and low mobility, GSS substantially increase the local supply of skilled women lowering wages locally and allowing the market to offer affordable education. These findings highlight the prominent role of women as teachers in facilitating educational access and resonate with similar historical evidence from developed economies. The students of today are the teachers of tomorrow.

This paper is a product of the Human Development and Public Services Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at jdas1@worldbank.org.

Students Today, Teachers Tomorrow? Identifying constraints on the provision of education

Tahir Andrabi Jishnu Das Asim Ijaz Khwaja*

JEL Codes: I21 (Analysis of Education), O12 (Microeconomic Analyses of Economic Development), J22 (Time Allocation and Labor Supply), L26 (Entrepreunership) and L32 (Public Enterprises; Public-Private Enterprises)

Keywords: Secondary Schooling, Private Schools, Teacher labor supply

*Andrabi: Pomona College; Das: Development Research Group, World Bank and Centre for Policy Research, New Delhi; Khwaja: Kennedy School of Government, Harvard University. Email: tandrabi@pomona.edu; jdas1@worldbank.org; akhwaja@hks.harvard.edu. This paper was funded through grants from the PSIA and KCP trust-funds and the South Asia Human Development Group at the World Bank. We thank Abhijit Banerjee, Esther Duflo, Karla Hoff, Rema Hanna, Caroline Hoxby, Hanan Jacoby, Brian Jacob, Ghazala Mansuri, Sendhil Mullainathan, Rohini Pande, Juan Saavedra, Tara Vishwanath, and seminar participants at BREAD (Yale), Lahore University of Management Sciences, LSE, NBER Education meetings, Harvard University, IUPUI, The World Bank, University of Michigan, University of Maryland, and Wharton for comments. We are grateful to Nirvana Abou-Gabal, Alexandra Cirone, Sean Lewis-Faupel, and Tristan Zajonc for research assistance. Assistance from the Project Monitoring and Implementation Unit in Lahore is also acknowledged. All errors are our own. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent.

1 Introduction

Despite the powerful global consensus created through the Millennium Development Goals, over a third of developing countries are "off-track" in achieving universal primary enrollment by 2015. One explanation for this poor performance is that the demand for education is inefficiently low. This is likely if parents do not fully internalize educational returns for their children.[1] In contrast to demand-based explanations, this paper evaluates the importance of a key supply-side constraint: the availability of affordable teachers. Teacher shortages can pose severe and persistent constraints. A high ratio of unskilled to skilled workers in the labor force implies a large skill premium, and thus, a high relative cost of training the uneducated. When credit markets are imperfect or long-term commitments are not credible, this high cost of training can lead to poverty traps (Ljungqvist 1993, Banerjee 2004).

The potential pool of teachers is limited in many parts of the developing world. Less than 12 percent of the population in Sub-Saharan Africa complete secondary education, with the more educated concentrated in urban areas. Educationists increasingly argue that there are severe teacher "shortages," a concern that resonates with the challenges faced in designing incentives for teachers to move to rural areas and to exert greater effort (UNESCO 2004, Urquiola and Vegas 2005, Chaudhury et al. 2006). Recent work on the decline in teaching quality in the United States also highlights the link between the supply of teachers and female labor force participation (Corcoran et al. 2004, Hoxby and Leigh 2004). Given this stress on teacher supply in low-income countries, it is therefore surprising that there is little micro-economic evidence relating a higher supply of potential teachers to better educational provision.

In this paper, we show that public investments in secondary education facilitate future

[1] This view has led to prescriptions such as conditional cash transfers whereby parents are incentivized to send their children to school (Schultz 2004, Filmer and Schady 2006). However, the high marginal cost of such programs may reduce their appeal: Estimates suggest that the cost per marginal child exceeds \$9,000 in Mexico and \$400 in Pakistan—figures that are very close to the GDP per capita of these countries (de Janvry and Sadoulet 2006, Chaudhury and Parjuli 2010).

educational provision by increasing the local pool of potential teachers and therefore decreasing the cost of providing education. In other words, the students of today become the teachers of tomorrow.

There are two steps to our argument. First, we show that the construction of government girls' secondary schools (henceforth GSS) in Pakistan had a large *causal* impact on the education market: Instrumental variable estimates suggest that villages where such schools were constructed are 27 percentage points, or three times more likely to see *private* schools emerge in the following years. The focus on private schools is important since the private sector better reflects local market conditions and thus aids in the identification of the teacher supply channel. [2]In the second step, we argue that GSS construction impacts private school location because it augments local teacher supply in an environment with low female geographical and occupational mobility.

The causal impact of GSS construction on private school location could also capture the effect of changes in demand: Educated mothers likely demand greater education for their children. In support of the "women as teachers" supply channel, we first document that: (a) private provision is affected *only* by GSS construction (girls' primary or boys' primary/secondary schools have little effect); (b) having a GSS more than doubles the number of women in the (median) village with secondary or higher education; and (c) the fraction of secondary educated females in a village has a large impact on private educational provision, while the fraction of similarly educated men does not. These facts could still be reconciled with demand-side explanations if the demand for education is primarily driven by mothers with secondary education (as opposed to mothers with primary education or fathers with any level of education). A more conclusive test is based on the effect of GSS construction on private school teachers' wages: Demand-side explanations suggest that teacher wages should

[2] The vast majority of private schools operate in a free and relatively unregulated market as for-profit, co-educational, English-medium schools that offer secular education (contrary to popular views, non-profit and religious schools play a small role in Pakistan, with at most a 3 percent enrollment share, Andrabi et al. 2006) and hire teachers from the local market. This is in contrast to the government sector where teacher hiring is governed by teachers' unions, state-wide hiring regulations, and non-transparent processes.

increase in villages with a GSS; supply-side explanations suggest the opposite. In support of the latter, we show that private school teachers' wages are 27 percent *lower* in villages with a GSS. With teacher wages accounting for close to 90 percent of the operational costs of private schools, this offers a substantial cost advantage. Moreover, consistent with the hypothesized mechanism, we find that this wage drop is higher in villages with more restricted female labor markets as proxied by village development indicators and sex-ratios.

To address the (potential) non-random placement of GSSs, we use an instrumental variables approach that exploits official eligibility guidelines for GSS construction from a Social Action Program in the 1980s. According to these guidelines, villages with higher populations were given a preference for GSS construction as long as there were no other GSSs (in neighboring villages) within a ten-kilometer radius. To operationalize the concept of neighboring villages in the absence of geo-referenced village locational data, we use the next highest administrative classification, the "Patwar-Circle" (PC), which typically covers four geographically contiguous villages and a land-area close to the ten-kilometer radius. Capturing the essence of this guideline, our binary instrument is an indicator for "local top-rank" that takes the value 1 if a village has the largest population among all the (neighboring) villages in its PC, and 0 otherwise.

Non-linearities and discontinuities in the eligibility rule (two villages with equal populations may or may not be eligible depending on their population rank within their neighbors) allow us to simultaneously control for polynomial effects of a village's own population, which have arguably independent effects on the educational market. Under the assumption that private school placement is not determined in the same non-linear and highly discontinuous fashion as the eligibility rule, the instrumental variables (IV) estimate is consistent.

The primary threat to this IV strategy is that unobservable attributes of villages with the highest population rank within a PC or the rank itself may be directly correlated with the existence of a private school. Specifically, estimates would be biased upwards if the government used the same strategy to allocate *other* public investments that may also directly impact the educational market, and/or if the private sector responds to rank conditions or factors associated with it in a similar fashion. We believe this to be unlikely, since the historical record shows that PCs are used only as revenue collection units, while political representation, and with it the delivery of public services, is centered at the Union-Council level, an alternate and non-overlapping classification.

Three empirical tests, in the spirit of Altonji et al. (2005), provide further support for the exclusion restriction. First, village socio-demographic characteristics are uncorrelated with the eligibility status of the village: Eligible and non-eligible villages do not differ along any observable dimensions other than those (population and area) on which the instrument is based. Second, the instrument does not predict the construction of any other type of public school nor any of a range of other public investments. Third, only *local* population rank corresponding to the ten-kilometer guideline matters. An analogous but more expanded local rank measure (top rank at the next higher administrative level which has a radius three times that of a PC's) does not predict girls secondary school construction. If entrepreneurs are more likely to pick locally top-ranked villages, we would expect this result to hold for the slightly more expanded top-rank measure as well. Furthermore, in a falsification exercise, we confirm that our instrument has no impact on private school location decisions in administrative units where there was no GSS construction. Thus, local population rank on its own does not affect private sector location decisions.

One natural question is whether this increase in the supply of teachers has led to an increase in educational provision or a sectoral shift from public to private schools. There are several reasons to think that the growth of private schools has had a positive impact on educational outcomes, both in terms of enrollment and learning outcomes. In a representative sample of households in the country (the Pakistan Integrated Household Survey 1998), overall enrollment is higher for villages with private schools (61 percent versus 46 percent), as is female enrollment (56 percent versus 35 percent). Moreover, Kim et al. (1999) provide strong causal evidence that private schools increase enrollment by showing that a randomly allocated subsidy for the creation of private schools in rural Pakistan led to increases of 14.6 and 22.1 percentage points in female enrollment for two of three program districts, likely by increasing school density (in a context where distance has important affects on enrollment). For the data used in this paper, enrollment rates in villages with private schools are 13 percentage points higher after conditioning on the presence of all types of public schooling, village population, and wealth, and accounting for all PC-level time-invariant factors.[3]

In addition, test scores of children in rural private schools are higher than those of their government counterparts even after accounting for possible child selection through IV and dynamic panel data methods. In tests administered to children in Grades 3 and 5, those in private schools outperformed public school students by 0.83 standard deviations in English, 0.67 standard deviations in the vernacular (Urdu), and 0.65 standard deviations in Mathematics (see Andrabi et al. forthcoming, Andrabi et al. 2011). This difference is further accentuated in cost terms because private schools are cheaper. The unionization and pay-grade of public teachers implies that the per-child costs of private schools are half those of public schools, a result consistent with findings from several countries around the world (Jimenez et al. 1991, Kim et al. 1999, Orazem 2000, Hoxby and Leigh 2004).

In thinking about the wider applicability of our results, it is worth separating the *existence* of supply-side constraints from their empirical identification. While such constraints are likely to affect educational provision more widely, there are several reasons why Pakistan is particularly well-suited for this empirical exercise. First, it has a large for-profit, unregulated private sector presence in education, accounting for 35 percent of primary school enrollment. This allows us to use variation in private sector provision of education as an indicator of variation in market forces. Second, government schools are segregated by both gender and level (primary or secondary), and labor markets are occupationally and geographically restricted for women. The combination of locally segmented markets for women

[3]While one may be tempted to instrument for private school existence in these regressions using the population rank instrument used in the paper (and we get even larger results if we do so), we do not believe the exclusion restriction is defensible in this case, i.e., top-ranked villages are both more likely to get a GSS and in turn a private school, and both these factors directly lead to increased enrollment.

with the gender and grade segregation of schools allows us to empirically isolate the impact of the local (gender and level-specific) supply shock on the private education market.

In environments with geographically integrated labor markets, the effect of an increase in local supply, while possibly just as important, would be harder to observe in the data since it would vary only at a higher level of geographical aggregation. Anecdotal evidence suggests that supply constraints in the form of teacher shortages are equally binding in Latin America or Sub-Saharan Africa. However, it may be harder to empirically demonstrate the effect of increasing local supply on the educational market if there are high migration rates.

Our results suggest that assuring a supply of teachers in rural areas of low-income countries is indeed a first-order problem that educational systems have to tackle. As in the United States (Rivkin et al. 2005), a consistent finding from observational and experimental studies in low-income countries is that augmenting teacher resources leads to better outcomes, whether through reducing class-sizes (Case and Deaton 1999, Urquiola 2006), reducing teacher absenteeism (Duflo et al. 2009), or providing additional teachers for poorly performing students (Banerjee et al. 2007). A natural question is whether *finding* these teachers in the first place is going to be a problem. The only randomized intervention (to our knowledge) that tried to increase the supply of schools through the private educational market failed precisely because teachers could not be found (Alderman et al. 2003).

The remainder of the paper is structured as follows: Section 2 is a brief guide to the institutional context and data. Section 3 presents the empirical methodology, and Section 4 the results. Section 5 concludes.

2 Institutional Background and Data

2.1 The Context

Pakistan, as in other South Asian and African countries, has experienced an explosion in the share of the private sector in education, both in terms of schooling availability and the enrollment share. The past two decades have seen more than a ten-fold increase in the number of private primary schools (3,800 in 1983 to 47,000 by 2005), and currently, over a third of primary-level enrollment is in the private sector, with the fastest growth coming from rural areas (Andrabi et al. 2008).[4] While this private school growth is impressive, it has generated more cross-sectional than time-series variation with growth mostly bunched in the 1990s. Hence, our paper exploits the cross-sectional variation in private school location to identify constraints to education provision. One of the key observations for the purposes of this paper is that since these private schools represent for-profit enterprises operating in a largely unrestricted market (there are no public subsidies and little regulation), their locational decisions are informative with respect to supply and demand factors in the educational market rather than public priorities or ideology (which may influence the location of public, NGO, or religious schools). Central to this argument is the importance of women (as teachers) in the provision of private education coupled with the limited availability of secondary-educated women in a restricted geographical labor market and the resulting impact on skilled female wages.

Key to understanding the private sector is its affordability and size. Andrabi et al. (2008) show that the median annual fee in a Pakistani rural private school in 2000 was Rs. 600, so that a month's fee was somewhat less than the daily wage rate of an unskilled worker.[5] The data show that there are few fixed costs in running a private school in Pakistan (private schools are often setup initially in the teacher/owner's house) with teachers' wages forming the bulk (90 percent) of the overall operational costs with typical schools utilizing four teachers and enrolling around 100 children. Moreover, most teachers in private schools are locally-resident females with (at least) a secondary education.

[4] Contrary to popular belief and media reporting, these changes have little to do with religious education. Andrabi et al. (2006) show that enrollment in religious schools, or madrassas is low (roughly 1 percent) and has remained constant since the mid-80s.

[5] In contrast, private schools (elementary and secondary) in the United States charged $3,524 in 1991. At 14 percent of GDP per capita, the relative cost of private schooling is 3.5 times higher in the US.

It is this reliance on female teachers that enables the private sector to offer affordable education.[6] In the context of a highly patriarchal society, limited geographical and occupational mobility for women implies that locally resident women offer a cheaper ("captured") supply of teachers. Female wages are indeed 30 percent lower than male wages after controlling for educational qualifications and experience (World Bank 2005). More than 70 percent of all women live in the village where they were born; less than 3 percent are engaged in off-farm work; and among those with secondary education and a wage-earning job, 87 percent are teachers or health workers. Safety concerns and a patriarchal society restrict the ability of women to find wage work outside the village where they live or in occupations other than teaching and publicly provided health care (World Bank 2005).

The presence of locally resident women can thus reduce the overall cost of wages for schools, but an assured supply in the *local* vicinity is critical. However, the supply of potential female teachers is low and varies across villages based on the availability of nearby schooling options. In 1981, there were *four* literate (adult) women (out of 242) in the median village in Punjab, the largest and most dynamic province in the country. Over 60 percent of villages in the province had *three* or less secondary-school educated women, and 41 percent had none. This was driven in part by a shortage of local secondary schooling options for rural women. A simple correlation in our regression sample between the availability of GSS and secondary educated women (in 1998) suggests that the presence of a GSS is associated with an increase of over 50 percent (compared to the median village without a GSS) in the (1998) percentage of women with a secondary education (from 3 to 4.6 percent).

These two features of the market for female skilled labor—low wages and limited supply—combined with the unrestricted and unsubsidized market for private schooling inform our empirical strategy. The presence of a GSS should generate cross-sectional variation in the availability of locally resident women with secondary education. If teacher supply constrains

[6] In comparison, wages for *public* sector teachers are five times higher for both men and women. As a result, per-child spending in rural private schools (Rs. 1012 annually) is half of that in rural public schools (Rs. 2039 annually), although available facilities are comparable across the two.

education provision (and there is limited mobility) this in turn should affect the likelihood of a private school existing in a village.

2.2 Data

We employ three data sources: (a) a complete census of private schools carried out by the Federal Bureau of Statistics in 2000; (b) administrative data on the location and date of construction of public schools in the Punjab province available from the province's Educational Management and Information Systems (EMIS 2001) augmented with the National Educational Census (NEC 2005); and (c) data on village-level demographics and educational profiles from the 1998 and the 1981 population censuses of Punjab, which provide both baseline and contemporaneous information on village-level characteristics.

We restrict our analysis to rural areas in the province of Punjab, the largest province in the country which hosts 60 percent of the population, two-thirds of whom live in rural areas.[7] Since the EMIS and the other datasets do not employ a common village coding scheme, we had to match villages in the different databases on the basis of their names. Using a combination of a phonetic algorithm and manual post-match, we were able to match over 90 percent of the villages across databases (23,064 of the 25,266 unique Punjabi villages in the 1981 census).[8]

In our final estimation sample, we restrict attention to villages that did not receive a girls' or boys' secondary school prior to 1981 and did not have such secondary schools in their neighboring villages. This reduces our sample to 9,333 villages, but affords two

[7] Not all data sets (e.g., EMIS, 1981 Census) were readily accessible for other provinces, and urban areas could not be matched at the granular level necessary to exploit the cross-sectional variation in private school location and GSS presence that we utilize in the paper.

[8] We also augment the public schooling data from the EMIS with more recent data from the 2005 National Educational Census. We are able to match some more villages using the NEC, but cannot use this as the primary data source because of insufficient information on the upgradation of schools from primary to secondary. Specifically, there is a chance in the NEC data that we incorrectly assign a village to have received a GSS by 2000 when it only had a primary school that was upgraded to a secondary school after 2000. Since our empirical strategy examines the relationship between pre-existing girls secondary schools and private schools (as of 2000, the date of the private school census), the correct thing would be to classify such a village as not having a GSS.

advantages. First, it allows for cleaner econometric identification and interpretation of the results as our instrument utilizes public school construction guidelines that were applied for GSSs constructed *after* the 1980s.[9] This also alleviates exclusion restriction concerns that arise if our instrument were to predict other public goods. Second, focusing on the shorter exposure (to GSS) periods is likely to better isolate supply-side effects since GSS construction probably impacts a range of demand factors over a longer time span. It is nevertheless reassuring to note that all of our main results hold in the full sample of villages, both in terms of statistical and economic significance, and several of these results are in fact stronger (Appendix Table II).[10]

Table I presents summary statistics for the final sample. Two and a half percent, or 232 villages, in this sample received a GSS between 1981 and 2001.[11] Conditional on existence,

[9] We are not aware of similar guidelines used in previous years. To the extent they were, we are reluctant to use the 1981 population (the earliest available census data at the village-level) to construct population rank for earlier years. Focusing on villages which did not receive secondary schools prior to 1981 also allows us to better control for village-level baseline data *prior* to the construction of a public school. For villages with pre-existing secondary schools, it is harder to discern whether differences in the baseline data arise from selection into villages or the exposure to the secondary school. While we could have also excluded villages which received girls'/boys' primary schools prior to the 1980s, this is too severe a restriction and would eliminate most of our sample. Finally, we are also concerned with pre-existing secondary schools since we believe they are more likely to reflect village wealth, size, or influence. We therefore also exclude villages whose neighbours' have pre-existing secondary schools, since that could have spillover effects through inter-village marriages and may mask supply-side channels. We are less worried about primary schools in neighboring villages affecting village demand, since there is considerable evidence that younger children do not travel outside their village to go to school (Alderman 1995, Andrabi et al. 2009).

[10] A couple of differences in the full sample results are worth noting and provide further support for our data restriction. Column (2) of Appendix Table II shows if we use the full sample, the instrument predicts boys' secondary school construction. This is not the case in the restricted sample (Column (5), Table II). The full sample association is not surprising because local rank criteria may have been used for BSS allocation in the past and/or the 1981 population may be an outcome of secondary school construction (since it is no longer a baseline variable as the full sample includes schools constructed before 1981). In addition, the impact of GSS on teacher wages in the village is noteworthy. While our restricted sample result shows that GSS presence leads to a lower wage (Table V), in the full sample we find that exposure to GSS has a non-linear effect on teacher wages. Initial exposure to GSS is indeed associated with lower wages, but prolonged exposure (more than 26 years) is associated with higher wages (the linear term on years exposure is negative while the quadratic terms is positive and of smaller magnitude). This is indeed consistent with our net supply impact interpretation of a GSS within the time-frame we are in (20 years) but suggests that, in the longer term, the demand effect may dominate: As more and more educated girls become mothers and grand-mothers, they impact educational demand. It therefore offers another important consideration for why restricting our analysis to the reduced sample is appropriate in identifying the (initially dominant) supply channel.

[11] This number is quite low relative to what the school construction guidelines would have suggested. While this is not surprising given that these guidelines were constrained by budgetary limitations, it may lead to concerns about the power of the instrument and the external validity of our results. We therefore address

the median age of a GSS is 14 years, therefore most were constructed early on in the 20-year period. There is a private school in one out of every eight villages, and the majority of these villages already had or received a primary public school. Finally, the number of women reporting secondary or higher education (eight or more years of schooling) increased from one in the median village in 1981 to nine by 1998.

3 Methodology and Empirical Framework

There are two broad empirical challenges that we seek to address in this paper. The first is to identify the causal impact of GSSs on subsequent private school existence. The second is to argue that this works, in part, through a teacher supply channel rather than an increase in the demand for education from secondary-educated women.

A simple framework outlines the private entrepreneur's problem, focusing on the role of the public sector and the econometric and interpretational issues in identifying the impact of a GSS on the educational market. An entrepreneur opens a school in village i if the net return, defined as the difference between total revenues and total costs, is positive.[12] For private schools in Pakistan, school fees and teachers' salaries account for 98.4 percent and 89 percent of total revenues and costs, respectively (Andrabi et al. 2008). We therefore approximate the net return for a school in village i as:

$$NetReturn_i = Fee_i * N_i - Wage_i * T_i \tag{1}$$

where Fee_i is the average private school fee for a single student, $Wage_i$ is the average private school teacher's salary, and N_i and T_i are the number of students enrolled and teachers employed. Since the schooling market may be geographically segregated, we allow wages and fees to differ across villages.

these in detail later in the paper.

[12]This assumes that there is no shortage of entrepreneurs (otherwise, not every positive NPV project will be undertaken). Incorporating such shortages does not change the qualitative results. The qualitative results also extend to a dynamic framework provided that the fixed costs of setting up schools is small.

The construction of a GSS increases the supply of teachers in the village, thus affecting $Wage_i$. However, it may also increase the potential demand for schooling, reflected in Fee_i. A reduced form expression for net return can then be written as:

$$NetReturn_i = \alpha + (\beta_1 + \gamma_1)GSS_i \ + \beta' X_i^D + \ \gamma' X_i^S \tag{2}$$

where X_i^D and X_i^S are village demographics and characteristics that respectively affect the demand for private schooling and the costs of running such schools. Variables in X_i^D and X_i^s include village population, measures of village wealth, adult literacy, and alternative schooling options. GSS construction has two effects in Equation (2): It alters the demand for private education by creating a more educated populace through β_1, and it affects the cost of setting up private schools by shifting the local supply of potential teachers through γ_1. We are interested both in the joint estimation of $(\beta_1 + \gamma_1)$ and in arguing that the there is a supply channel (i.e., γ_1 is positive and significant).

Since the net return a private school earns is not observed, we treat net return in Equation (2) as a latent variable in a probability model such that $Prob(PrivateSchoolExists) = Prob(NetReturn_i > 0)$, and estimate a version of Equation (2):

$$Private_{it} = \alpha + (\beta_1 + \gamma_1)GSS_{it} + \beta' X_{it} + \ \sum_r \gamma_r' S_{irt} + (v_i + \varepsilon_{it}) \tag{3}$$

where $Private_{it}$ is a binary variable that takes the value 1 if a private school exists in village i in time t and GSS_{it} is a binary variable that takes the value 1 if a GSS exists in village i at time t. X_{it}^D observed characteristics village characteristics at time t. S_{irt} are other government schooling options (primary boys/girls schools and boys secondary school) at time t, where each option is indexed by r. The error term, $(v_i + \varepsilon_{it})$, consists of a time-invariant unobserved component, v_i, and a random component, ε_{it}. The main identification challenge is that the presence of a GSS in village i in time period t is likely a function of the latent unobserved components of the village/region:

$$GSS_{it} = \alpha' + \varphi X_{it} + (\lambda_i + \mu_{it}). \tag{4}$$

Thus, the OLS estimate of $(\beta_1+\gamma_1)$ in Equation (3) is biased and inconsistent if $cov(\nu_i, \lambda_i) \neq 0$. While first differencing Equation (3) helps, the estimated $(\beta_1 + \gamma_1)$ in such a specification would still be biased if $cov(\varepsilon_{it}, \mu_{it}) \neq 0$ (i.e., there are time-varying covariates that determine receiving a GSS and affect private school presence). Therefore, we instrument for GSS construction using program guidelines for a school expansion program undertaken in the 1980s.

3.1 Identification Strategy

Our instrumental variables strategy exploits the fact that the regressor of interest, the construction of a GSS, is partly based on a deterministic function of a known covariate, village population. If this deterministic function is non-linear and non-monotonic, it can be used as an instrument while directly controlling for linear and polynomial functions of the underlying covariate itself (see Campbell [1969], and Angrist and Lavy [1999]).

GSS construction after 1981 was a consequence of the 1980 Pakistan Social Action Program (SAP). Specific guidelines affected where these schools could be built. In particular, the recommended guidelines for opening a new GSS specified a preference for higher village populations and stipulated that there be no other GSS within a ten-kilometer radius.

In order to capture this guideline, we construct a binary assignment rule, $Rule_i$, that takes the value 1 if the village is the largest village (in terms of population) amongst nearby villages and 0 otherwise. This captures the radius criteria. If a village is not the largest village amongst its neighbors, the neighbor would receive a GSS first given the stated preference for population. Provided this school is near enough, the village will be less likely to receive its own public school.[13] In the absence of precise village location data, we use the next

[13] Another alternative is to use the radius-rule directly and assign $Rule_i = 0$ if there is a village in the

highest administrative classification, the "Patwar-Circle" (PC), which typically covers four villages, to approximate the radius rule. In terms of actual land area, this is a reasonable approximation; dividing the size of the province by the number of PCs shows that one school in every PC would satisfy the radius requirements of the rule. Formally:

$$Rule_i = \begin{cases} 1 \; if \;\; Population_i = \max\limits_{j \in PC_i}(Population_j) \\ 0 \; if \;\; Population_i < \max\limits_{j \in PC_i}(Population_j) \end{cases}$$

Since GSSs could have been built in any year between 1981 and 1998, we assign a value of one to $Rule_i$ if it was the largest village in its PC based either on its 1981 or 1998 population. In addition, for the 4.5 percent of villages in our sample that are alone in their PC, we assign a value of 0 to the instrument. Our results are robust to the using either 1981 or 1998 population exclusively or assigning the value 1 to $Rule_i$ for single-village PCs.

The eligibility rule is non-linear and non-monotonic in population. It drops to 0 for larger villages when there is an even larger neighboring village within the PC. In using this rule as an instrument, we are thus able to explicitly control for continuous functions of village population (these covariates have a large direct impact on the existence of a private school). We also include a full set of PC fixed effects in our specification, thus exploiting rank variation only within a small set of proximate villages. Our final specification is of the form:

$$Private_i = \alpha_{PC_i} + (\beta_1 + \gamma_1)GSS_{it} + \beta_1' Pop_{i81} + \beta_2' Pop_{i81}^2 + \beta_3' Pop_{i98} + \beta_4' Pop_{i98}^2 + \beta' X_{it} + \sum \gamma_r' S_{irt} + (v_i + \varepsilon_{it}) \quad (5)$$

where the X_{it} controls also include indicators of village wealth and area. We estimate Equation (5) using $Rule_i$ as an instrument for GSS_{it}.

patwar-circle that has a GSS. This is problematic since we are worried about the endogenous placement of GSS in the first place.

With PC fixed effects and population controls, the remaining variation that the rule exploits is likely uncorrelated with the demand for private schooling. Nevertheless, there may still be concerns that the same local rank criteria is relevant for the provision of other public goods. In Section 4, we present several robustness tests to check for the validity of the exclusion restriction. Specifically, we show (a) that our instrument does not predict the construction of other public goods and (b) that it is the local (within-PC) population rank that matters rather than a village's population rank in the next larger administrative unit above a PC, where the radius rule would less likely apply.

3.2 Isolating the Supply-Side

To separate supply from demand-side channels, we propose two strategies based on the relative effect of educated women versus educated men in the location decisions of private schools (the quantity margin) and the costs of operating private schools in villages with and without a GSS (the price margin).

On the quantity margin, a supply-side channel suggests several patterns. In particular, we expect that: (a) since most teachers in private schools report at least a secondary education (98 percent), *secondary* schools should have a larger impact on private school existence than primary schools; (b) the effect of GSS should be larger than that of boys' secondary schools; (c) villages with a GSS should report a larger stock of educated women;[14] and (d) private school existence should respond more to women with higher education than men with higher education. While results in the expected direction lend support to the supply-side channel, explanations based on the relative importance of women versus men or secondary versus primary education in fostering the demand for education cannot be ruled out.

More conclusive evidence for the presence of the supply-side channel comes from the price margin. If private schools locate in villages with a GSS due to increases in demand,

[14] This test is relies on there being limited migration. To the extent that educated women migrate out (in), the estimates could be attenuated (overestimated). With female migration rates around 15% (Hamid 2010) we don't perceive this as a substantial concern.

we should see higher teachers' wages in such villages. Conversely, if the GSS effect works through the supply channel, we should observe lower wages. Therefore, one should test for differences in skilled women's wages in villages with and without a GSS.[15]

However, the challenge in doing so is a data issue: The only available village-level data that captures skilled women's wages is the private school census, which records average teacher wages in all private schools.[16] Since we do not observe wages in villages without private schools, a simple correlation of wages and GSS may be biased, with the bias depending both on how GSS were placed and on the truncation of the wage distribution due to missing wages in villages without private schools. We follow two approaches to address the selection problem. We use a Heckman selection model, where the selection stage is the probability of observing a positive wage, which corresponds to having a private school in the village. Another alternative is to use the "control-function" approach, where we condition on the predicted probability of observing a non-missing value of the wage-bill in the wage equation (Angrist 1995). Details of both approaches are in Appendix I.

We should caution that we cannot structurally estimate the size of the supply-side effects. For instance, simultaneous changes in the demand for schooling due to GSS construction imply that the supply-side impact of GSS construction on (decreasing) the wage-bill represents a lower bound. Therefore, our strategy indicates the presence of a supply-side impact but has less to say about its size.

[15]If there is a preference to teach in private schools, increased demand could drop wages as teachers may be willing to accept lower wages in new private schools. However, instead there is a strong preference for *public* schools (better pay and easier job). In addition, the labor market for public and private schools is quite different, with the former being non-local and the latter local. Moreover, within private schools, the market is not stratified so it is unlikely that there would be systematic compensating differentials across different private schools. One may also be concerned about whether private school wages are meaningful if the owners also teach (wages may be confounded with profits). We do not think this is a substantive issue. A detailed examination of profits using the smaller sample in the LEAPS database, suggests that median profits are quite comparable to a teacher's wage. Moreover, most of these schools do employ non-family/paid labor and therefore reported wages indeed reflect the opportunity cost of hiring (local) skilled women,

[16]An alternate data source is the Pakistan Integrated Household Survey (PIHS). Unfortunately, given the small number of villages that received a GSS, the available sample sizes are too small in the PIHS. With the sample restrictions in our paper, we find only three villages in the treatment and thirty-one villages in the control set for these data. Moreover, since the majority of (the few) women who work in non-farm activities are teachers, and the vast majority of private school teachers are women, the private school wage bill is likely to reflect the wages of skilled women.

4 Results

4.1 Instrumental Variable Strategy: First Stage and Specification Checks

To clarify the identifying assumptions needed for our IV strategy, Figure I illustrates how the existence of private schools and the binary instrument covary with the 1981 village population (the relationship with 1998 population is similar). Here, we plot $Rule_i$ for all villages in our sample and the *non-parametric* relationship between private school location and village population. There are both "eligible" ($Rule_i = 1$) and "ineligible" ($Rule_i = 0$) villages *at all population levels*. We can thus compare two villages with the same population, one of which was eligible to receive the GSS and another that was not, allowing us to exclude the direct effect of population on private school existence. Further, the non-parametric relationship between private school existence and village population is approximately linear; it is therefore likely that linear and quadratic population terms in the regression specification sufficiently control for the underlying relationship between village population and private school existence.

Table II, Columns (1) and (2) present regression estimates using the eligibility rule as a predictor for the location of GSS. Column (1) runs a probit specification with linear and quadratic controls for population, and shows that an eligible village was 1.24 percentage points more likely to receive a GSS. Column (2) augments the first stage with other village-level public goods and PC fixed effects, resulting in similar point estimates that are significant at the 1 percent level of confidence: Villages with $Rule_i = 1$ were 1.6 percentage points more likely to receive a GSS. Although the point estimate seems small, this is because few girls' secondary schools were constructed. In fact, this estimate represents an almost 100 percent increase over the fraction of ineligible (instrument = 0) villages that had received a GSS by 2001. In addition, both the basic and the more demanding first stage are at or above the proposed critical thresholds for detecting weak instruments (Stock et al. 2002).

Instrument Variables Strategy: Exclusion Restriction

To assess the validity of the exclusion restriction, we first confirm that there are no statistically significant baseline differences in educational levels for women or men nor in their age distribution between eligible (instrument = 1) and ineligible (instrument = 0) villages (Appendix Table I). The only differences are in the initial population and area, which arise directly from the construction of the instrument and are controlled for in the IV specifications. Moreover, there are no differences in 1998 in other village socio-economic attributes such as the extent of permanent housing, media access (TV and radio), men/women with national identification cards, or sex-ratios. This is reassuring since it is consistent with random assignment of the eligibility rule across villages.

The exclusion restriction could also fail if the government used the same village population-rank criteria for allocating other investments. Of note is that PC is a historical land revenue recording unit and has never been used as a jurisdiction for policy making purposes such as the delivery of public services or political representation. The smallest administrative political unit has always been the somewhat larger Union Council (UC), with little overlap between the two. Columns (3) through (8) in Table II directly assess this by demonstrating that our instrument, local rank in a PC, does not predict any other government investments apart from GSS. Columns (3) to (5) respectively show that local rank does not predict girls primary or boys primary/secondary school's placement. While the point estimates for primary schools appear similar to that of the GSS, they represent less than a 2 percent increase relative to the comparison group (i.e. over 50% of ineligible villages also had a primary school by 2001) as compared to the 100 percent increase for GSSs between eligible and ineligible villages. Columns (6) through (8) consider other public goods, such as access to potable water, electrification, and permanent housing structures, and again find no evidence that publicly provided goods are higher in eligible villages.

A third possible concern is that being a top-ranked village in a region is important in itself and that our instrument does not reflect the ten-kilometer-radius rule but a more general rank effect. For example, one may posit that private entrepreneurs also choose the largest village within a PC. While we believe such a concern is less plausible (private school entrepreneurs are almost always local to the village, with schools typically setup in the entrepreneur's house), one can test the (independent) importance of local rank by checking if the local rank within the next largest administrative unit after the PC, a Qannongoh Halqa (QH), predicts GSS placement. There are roughly ten PCs in a QH, and hence, the radius rule is unlikely to apply within a QH (villages are a lot further than ten kilometers apart in a QH). However, if local rank is important in general, one would still expect that being the top-ranked village in a QH would predict having a GSS. Column (9) shows that being the top-ranked village in the QH does not predict GSS placement. Column (10) adds our instrument, local rank in the PC, and shows that our instrument still predicts GSS placement while the analogous local rank measure at a larger geographical level (the QH) does not. This lends further support that our instrument predicts GSS placement because of the ten-kilometer-radius rule rather than some inherent characteristic about top-ranked villages within administrative units. Moreover, as we detail in the next section, PC-rank only matters in regions where we would expect it to (i.e., where a GSS was provided). [17]

4.2 GSS Impact on Private Schools

Table III first presents OLS results based on Equation(5).[18] The construction of a GSS

[17] In addition to these checks, we also conducted a placebo experiment. Starting from the full sample, we randomly grouped villages into "fake" PCs with four villages in each PC and classified villages as eligible using the new PC classifications and their actual 1981/1998 populations. We then estimated the reduced form relationship, $cov(Private_{it}, GSS_{it}|Pop)$. These steps were repeated five thousand times to generate a distribution of estimated coefficients under random assignment of villages to PCs. Our actual reduced form coefficient lies within the top 1 percentile of the distribution of reduced form coefficients generated by the fake PC simulations (the mean and median for the fake distribution are essentially zero). In other words, it is extremely unlikely that the coefficient we obtain is an artifact of a village being large; what matters is the specific assignment of villages to PCs.

[18] We focus on the existence of private schools rather than their enrollment share. Most variations in the number of children enrolled in private schools is driven by the extensive (whether or not there is a private school in the village) rather than the intensive (variation in private school enrollment conditional on existence) margin. Our results are similar if we look at private school enrollment. We prefer the extensive margin since the data on enrollment is noisier.

increases the probability of a private school in the village by 9.5 percentage points [Column (1)]. An equally significant determinant of private school existence is village population; the GSS effect is similar in magnitude to increasing (1998) village population by around 1500 individuals (slightly below a standard-deviation increase). Note that the specification includes a full set of village-level controls, including exposure to other types of public schools and PC fixed effects. Column (2) addresses any selection concerns arising from time-invariant village effects by first-differencing (1998 less 1981 values) the data at the village level. The effect of receiving a GSS on change in private school existence increases slightly to 9.7 percentage points. Propensity score estimates also yield similar results: A GSS increases private school existence probabilities by around 10 percentage points, depending on whether we use local linear regression or kernel matching (results available with authors). For the sake of comparability we use the same baseline year to difference the dependent variable (i.e. it takes the value one if the private school was created after 1981). There is a concern that this may be too soon and private schools made before 1984 should be excluded (giving at least three years past primary for the GSS to produce potential teachers). However, since most (99%) private schools were created after 1984 in our sample, doing so does not qualitatively affect our results and so we stick to 1981 as the baseline year for all variables in the first-differenced specification.

Figure II provides a simple illustration of our instrumental variable estimates by dividing villages into four population quartiles, averaged over 1981 and 1998 populations. The top panel compares the percentage of villages with a GSS in the "eligible" ($Rule_i = 1$) group compared to "ineligible" ($Rule_i = 0$) group. Over the entire sample, this difference represents the first-stage of the instrumental variables (IV) estimate, $cov(GSS_{it}, Rule_i)$. The bottom panel then compares, over the same population quartiles, the percentage of villages with a private school in the "eligible" and "ineligible" groups; this is the reduced form for the IV estimate. The figure shows that the instrument varies in every population quartile so that our results are not driven by variation in a single population group. For all population quartiles, the first-stage indicates that eligible villages were more likely to receive a GSS. In addition, the reduced form suggests that, controlling for population, villages that were eligible to receive a GSS were also more likely to see private schools arise at a later date.

Columns (3) to (5) of Table III present the corresponding IV regression coefficients. In Column (3), we present estimates using a linear IV specification. Given that both the existence of a GSS and the presence of a private school are binary variables, Columns (4) and (5) present estimates of the Average Treatment Effect (ATE) and Treatment on Treated (ATT) using a bivariate probit specification.

Column (3) shows that the estimated coefficient of GSS on private school existence increases from the OLS and first difference specifications to 1.50 in the linear IV specification, and the significance drops to the 10 percent level. Columns (4) and (5) implement the bivariate probit specification and report analytical standard-errors computed using the delta method. The point-estimate from the bivariate probit is still large but less than a fifth that of the linear IV and significant at the 10 percent level of confidence for the ATE and the 1 percent level for the ATT. The biprobit estimates suggest that private schools are 25 to 27 percentage points more likely to locate in villages with a GSS—a more than 200 percent increase over the comparison group (villages without a GSS) probability of 12.3 percent. The linear IV estimates are larger, and it is likely that the structure of the data accounts for this difference. As shown in Chiburish et al. (2010), the confidence intervals obtained from linear IV estimates are particularly large when treatment probabilities are low and the model includes additional covariates. Both of these problems are salient in our context: Given budget constraints under SAP, only 2.5 percent of the sample actually received the treatment, and for the exclusion restriction to hold, linear and quadratic population terms must be included in the specification (see Chiburish et al. [2010] and Appendix II). As such, our preferred estimates are from the bivariate probit specification.

The larger IV estimates suggest that time-varying omitted variables that increase the likelihood of private schools are in fact negatively correlated with GSS construction. There are several reasons why one may plausibly expect this. One interpretation made by Pitt et al. (1995) in the case of Indonesia is that governments act altruistically, trying to equalize differences between villages. Villages with lower responsiveness of demand to school construction received GSSs, and these were also the villages where private schools were less likely to locate. However, the Pakistani context suggests additional explanations, as well. Schools are often also targeted to villages with powerful/feudal local landlords and officials. These are precisely the villages where the demand for education is likely lower and less likely to increase over time. Moreover, given the requirement to give land for free for school construction, these schools were constructed in areas where land prices were also low. To the extent that low land prices are associated with poor educational returns, we would expect similar results to those documented here.

A Further Check of the Exclusion Restriction

Columns (6) and (7) present an additional check for our instrument by showing that the reduced form only holds where one would expect (i.e., regions where at least a GSS was provided). Here, we divide villages into two sub-groups, "program regions," where at least one village in a broadly defined area (we use QH, the unit larger than a PC) received a GSS and "non-program regions," where no village in the QH received a GSS. Note, in particular, that even if we do not know how *regions* were selected, comparisons across program and non-program areas are instructive. In particular, if population rank within the PC has no independent effect on the probability of setting up a private school, we should find a strong relationship between private school existence and eligibility for villages in program regions but *not* in non-program regions. A contrary result in non-program areas would suggest a violation of the exclusion restriction in our IV strategy. Our results confirm that population rank with the PC has an effect on private school location only in program areas, providing further support for the instrument. Column (6) shows that for program regions, eligibility increases the probability of a private school by 3.8 percentage points; conversely, in non-program regions, eligibility has no impact on private school existence [Column (7)].[19]

4.3 Potential Channels: Evidence for Supply-Side Effects

We now consider whether the causal impact of GSS on the educational market works through a supply-side channel. As described in Section 3, we do so by examining the impact of GSS on both the quantity and price margins.

Quantity Margin

If private schools arise because of the availability of "women as teachers," we expect a GSS to have a larger impact relative to other types of public schooling. Columns (1) and (2) in Table IV present estimates from a linear probability model and a first difference specification, both of which include PC-level location dummies. Both specifications confirm the importance of GSS relative to other types of public schooling. Column (1) shows that the coefficient for years of exposure to a GSS is almost four times as large as that of the next most important public school type. The first-difference specification shows that by better addressing time-invariant village selection factors the importance of GSS is further magnified: The change (from 1981 to 1998) in whether a village has a GSS or not is the *only* schooling variable that matters, and the magnitude of the effect is large. In contrast, whether a village received a boys' primary/secondary or girls' primary school between 1981 and 1998 has no affect on the likelihood of a private school setting up in the village (in fact, there is a negative association for boys' primary schools).

Columns (3) to (6) present the next logical step. We assess the correlation between educated women and the presence of a GSS for a variety of specifications. In both the OLS and first-difference specifications, a GSS increases the number of adult women with higher

[19] We also estimated a single pooled specification that controls for potential differences between program and non-program regions by including the predicted propensity (and its quadratic) of being a program region. Results (not shown) were very similar; the coefficient of the interaction between GSS and a program region is large and highly significant. In contrast, the eligibility rule in non-program regions has no effect on private school placement. Replicating the first-stage, linear IV, and biprobit estimates for program regions also produces similar results and with more statistical significance given a stronger first-stage (not surprising, since identification is achieved only off the variation in program regions).

levels of education (equal to eight or more years of schooling) by 9.5 to 10.8 more women, and the estimated increases are significant at the 1 percent level of confidence. Although this appears to be a small effect, it represents a substantial change in the *stock* of educated women. With eight women in the median village (without a GSS) in 1998 reporting higher levels of education, a GSS more than doubles this number. Column (5) utilizes a similar IV strategy and, as before, shows that while the IV estimate is significant, it is substantially larger than the OLS estimate. This is due to the relatively small first stage coefficient (see Table II). Column (6) makes this clear by presenting the reduced form estimate. While the large magnitude of the IV estimate is difficult to take literally and we believe the OLS/first difference estimates are more realistic, the point is that GSS existence substantially increases the number of educated women in the village even when potential selection concerns are taken into account.[20]

Columns (7) and (8) then examine the importance of secondary school educated women for the existence of a private school. In both the OLS and first-difference specifications, the impact of women with eight or more years of schooling is large and very significant, while the percentage of similarly educated males has *no* impact on the existence of a private school. In fact, the point estimate is one-tenth that of the female effect (and of the wrong sign in the first-difference specification).

Another potential approach to isolating the supply-side is to use variation in the timing of the public school construction since supply-side channels suggest that private schools will emerge five to eight years after the construction of a GSS (or three years if there was a preexisting primary school). Unfortunately, the data are too limited to exploit this variation but there is suggestive evidence that this is indeed the case.[21]

[20]We should note that the OLS/first-difference are large enough to generate (the few) teachers one would need for the supply channel, but not enough to produce sufficient educated mothers that one would expect if the demand channel were the primary driver. While the IV estimates could generate such a demand channel, they are implausibly large: The median village in our sample has only 9 women with higher education in 1998, with a mean of 26 and, with a typical GSS only graduating around 5 or so girls per year. Even by 2005, an increase of 220 women is therefore quite implausible.

[21]We require villages with both private schools and a GSS. Since only 232 villages received a GSS, and of these, 26 percent had a private school, we are unable to identify any discontinuities using the 60 or so villages

While these results by themselves may not rule out a demand-side channel, they do substantially constrain the routes through which it can work. Fathers' education could not stimulate demand for children's education (since boys' schools have no effect); primary schooling for mothers could not be enough to stimulate demand; mothers' schooling must therefore have a non-linear effect on the demand for children's education.

Price Margin

Table V provides further evidence for a supply channel by examining the price margin. Recall, in sharp contrast to a demand-channel, a supply channel would suggest that GSS construction would lead to a *fall* in private school teacher (i.e. skilled women) wages. We compare the average (log) teacher salary in private schools in villages with and without a GSS using data from the private school census. Column (1) presents the OLS results in the sample of villages for which we have teacher wage data. We include PC FEs in all specifications. The results are large and significant: Private schools in villages with a GSS report a 27 percent lower average (teaching) wage.

Columns (2) through (5) correct for selection into the wage sample. Columns (2) and (3) present results using Heckman's selection model, and Columns (4) and (5) use the "control function" approach (see Appendix I). In both approaches, identification is based on the non-linearity of the selection equation (see Duflo [2001] as an example). Augmenting the instrument set with potential candidates that are correlated to the probability of having a private school but uncorrelated to the wage-bill can help with the identification and the efficiency of the estimator. Following Downes and Greenstein (1996), we propose using the

that have both. An alternate strategy is to check whether there is a difference in the existence of a private school based on years of exposure to a GSS. Here, we do find some suggestive evidence. In particular, private schools exist in 22 percent of villages with 15 years or less of exposure to a GSS , and in 33 percent of those with more that 15 years. Moreover, it is really only older GSSs' which have an impact. We can conduct a similar placebo exercise as in Table III, Column (7) except we now include villages in the "non-program" group if they or a village in their QH received a GSS less than 5 years ago. Similar to the Column (7) result, we find that there is no reduced form effect of the instrument in this sample, i.e. it is only 5 years or more exposure to GSSs that matters. Finally, consistent with the supply channel, we find from a smaller but more in-depth sample that the female private school teachers are in the age-group that would be consistent with the GSS construction period - the median private school female teacher age is 22 with over 90 percent between 18-32 years of age.

number of public boys' primary schools as an additional instrument in the selection equation. In the presence of competitive schooling effects, private schools should be less likely to set up in villages where there are public boys' primary schools. Additionally, such schools are unlikely to affect the wage-bill of the entrepreneur directly since public school teachers are rarely, if ever, hired locally and because their wages are fixed and centrally determined. While we remain cautious in pushing this instrument since primary schools for boys may be endogenously placed, it does serve as a robustness check on the identification based on non-linearities in the selection equation. Columns (2) and (4) use the functional form of the selection equation to achieve identification, and Columns (3) and (5) introduce the additional instrument. The results are similar to the OLS estimates, with estimates of 27 to 28 percent[22] suggesting that selection into the non-zero wage sample is of limited importance.

Columns (6) and (7) present tentative evidence that wage declines due to a GSS are larger in villages where labor markets for women are more restricted and localized, i.e., the interaction terms of GSS existence and the village progressivity indicator are positive. In Column (6), we look at the differential effect of GSS construction on wages for more and less progressive villages using the female/male ratio for children under the age of 14 as an indicator of progressivity/gender bias. Arguably, villages with a lower female/male child ratio may be more conservative with fewer labor market opportunities for women outside the immediate vicinity of the house. Indeed, villages at the 25th percentile of the distribution (female/male ratio of 0.86) see a wage decline of 58 percent due to GSS construction, compared to essentially no decline for villages at the 75th percentile of the distribution (female/male ratio of almost 1).

In Column (7), we look at analogous results using households per capita with access to radios as an indicator of village-level development. While the results for the interaction term are only significant at the 26 percent level in this case, the signs are in the expected

[22] Since our dependent variable is log (wage), the coefficients of about -0.32 on the GSS existence dummy represent a decrease of approximately 27 percent in average wages. For example, in Column (2), the coefficient implies that, in villages with a GSS, wages change by a factor of $e^{-0.3207}$(or 0.7256) , which is equivalent to a 27.44 percent decline.

direction. Wages decline by 46 percent decline in villages where no houses have access to radios (6 percent of the sample), compared to a 26 percent decline in villages which are at the 75th percentile of the radio access distribution. While encouraging, these results are at best tentative. Endogenous variation (these variables are only available in the 1998 and not baseline, i.e. 1981, census), as well as the suitability of these two indicators as proxies for the restrictiveness of the female labor market, requires that they be viewed with some caution.

One may posit more nuanced demand-side explanations for such wage effects that introduce heterogeneity in the quality of teachers. We believe such stories are neither plausible nor empirically supported. For example, if increased demand spurs perverse competition across (private) schools (with parents unable to judge/evaluate quality), this may result in a "race to the bottom." In such a story, wages drop in villages with a GSS not because of the supply shifter but because the increased demand causes so much school entry/expansion that teacher quality (and hence wage) drops. However, given the large average wage drops we find, this would imply that the quality of the marginal teacher is substantially worse. Yet, not only is this implausible since parents are reasonably aware of teacher quality (Andrabi et al. 2009) but our regressions control for the number of schools and show that villages with more schools have *higher* wages. In other words, competition raises, not perversely lowers, wages.

4.4 Discussion

The wage estimates we obtain are also broadly consistent with a set of arbitrage conditions that should hold in equilibrium under a supply-side explanation. To see this, consider an entrepreneur who plans to set up a private school in a village *without* a GSS. She has several potential options, and for our results to be plausible, it must be (as we argue below) that these options are not viable.

First, she could hire a male instead of a female teacher. If we assume that men have fewer/no occupational and geographic mobility restrictions, this suggests that (equivalent) men must command at least 27% (the GSS impact on teacher wages) higher wages then women. If they didn't, then private schools could setup in villages without a GSS by hiring (local/non-local) men rather than women as teachers. Andrabi et al. (2008) show that men (with the same observed characteristics) indeed earn 33 percent more than women, suggesting that men do not offer a viable teaching alternative.

Second, the entrepreneur could try to setup a school with a larger initial class-size in order to pay for the greater cost of hiring a male teacher. However, this has the trade-off of lowering quality. A natural constraint here is that student performance in the private schools must exceed that in the (free) public schools. Andrabi et al. (forthcoming) uses GMM methods together with children who switch school types to show that the yearly value-added of private schooling is around 0.25 standard-deviations. Although the estimates from the experimental literature on class-size reductions vary somewhat, a number of studies suggest gains of 0.2 to 0.3 standard deviations due to a reduction of four to ten students (Angrist and Lavy 1999, Krueger 1999, Muralidharan and Sundararaman forthcoming). Given median wages and school fees in Punjab, to generate enough revenue to cover the 33 percent higher wages of a male teacher, the school would need a class size that is seven children more than the median private school. This suggests that the quality drop from the increase in class size required to hire male teachers would almost entirely offset the private school advantage in these villages and hence not be viable. In other words, parents would choose public schools instead if the private school had a larger class size.

Third, the entrepreneur could set fees 33 percent higher than the current levels. Using data from Pakistan, Carnero et al. (2010) structurally estimate the elasticity of private school market shares to fees and find that a 1 percent increase in prices reduces the market share per private school by 1.2 percent. Given this high price elasticity, private schools would therefore not be able to increase profits by raising fees, ruling out this arbitrage opportunity, as well.

5 Conclusion

Achieving universal primary education remains an elusive goal in many developing countries. While governments can choose to invest greater amounts in providing and subsidizing the costs of public schooling, the budgetary implications of such a task are daunting. Private educational provision is an increasing presence, particularly in developing countries, with shares exceeding 20 percent at the primary level in a large number of countries.

The crucial question is whether the market can offer affordable and quality education at a scale that can complement the public sector in achieving universal enrollment goals. This paper underscores that for this to happen local supply-side constraints need to be alleviated. While not surprising at the aggregate level, the result that (teacher) supply curves are not perfectly elastic at the *village* level can generate poverty traps in credit-constrained environments (Ljungqvist 1993 and Banerjee 2004).[23] Higher returns to education may perversely lead to *declines* in the provision of education if the returns increase as a consequence of higher wages in non-teaching professions. Moreover, locally upward-sloping supply curves have consequences for the pricing of voucher schemes. Depending on the elasticity of supply, increases in demand through vouchers may lead to simultaneous increases in prices, a decline in quality (in price-capped schemes), or both.

In contrast to calls for larger primary school investments at the expense of secondary schools, our findings suggest that both play a role. Public investments in secondary schools increase the supply of potential teachers locally and foster the growth of private schools, potentially leading to a virtuous cycle of human capital accumulation.

The changes documented in this paper represent more than just a sectoral realignment from public to private schools. Work in Pakistan and other countries suggests that the

[23] An upward sloping supply curve at the local level reflects supply constraints in the educational sector as it arises due to local labor market restrictions. There is a natural parallel with the literature on credit constraints. Evidence for such constraints is whether the cost of borrowing increases with the amount for *individual* firms. Again, that the cost of borrowing increases with the amount at the *aggregate* level is not surprising; conversely, firm-specific borrowing costs that increase with the amount borrowed lead to several important policy conclusions (see Banerjee and Duflo 2004).

growth of private schools represents an improvement in overall education, both in terms of raising educational quality and by allowing for higher overall and female enrollment in the village by reducing the distance to school and increasing the density of schooling options. As in other low-income countries, private schools appear to offer higher-quality education at far lower costs. The unionization and pay-grade of public teachers implies that per-child costs of private schools is half that of public schools (Jimenez et al. 1991, Kim et al. 1999, Alderman et al. 2003, Hoxby and Leigh 2004).

The importance of supply-side constraints however, cautions against over optimism regarding market educational provision and emphasizes the public sector's role. This is particularly important given a new round of pessimism about public sector provision. In South Asia for instance, the public sector is widely regarded as broken. With teacher absenteeism exceeding 40 percent in some areas (Chaudhury et al. 2006) and political imperatives making reform difficult (see, Grindle 2004), the private sector is increasingly viewed as a viable alternative (Tooley 2005, Tooley and Dixon 2005).

This paper shows that private sector schools do not arise in a vacuum. Previous public investments "crowd-in" the private sector so that government schools are not only contemporaneous substitutes but also temporal complements with private sector provision (Tilak and Sudarshan 2001 confirm a similar complementary relation in India). Moreover, analogous supply-side constraints likely exist at higher education levels. The fact that the private sector hasn't made as much in-roads in secondary schooling suggests that teaching supply constraints have yet to be alleviated at that level.

The public sector is then left with a tricky task in these environments. If the private sector is indeed to play a role in educational provision, initial investments from the public sector are required to build up the necessary supply of teachers. However, once the private sector enters the local market, the public sector becomes a direct competitor for teachers in a very limited market. Since public school teachers are paid substantially more than their private sector counterparts (five times more in the case of Pakistan), this direct competition coupled with poor accountability in the government sector now hurts educational provision. If, as we suggest, private schools represent an increase in the quality of education and raise overall enrollment levels rather than a shift in its sectoral composition, the public sector has to do enough, but not too much.

Appendix I

Selection Issues in the Wage Bill

Since we only observe the wage bill in villages where there is a private school, a concern described in the main text is that simple OLS estimates may be biased if such selection is not accounted for. Here, we provide details on two approaches that we use in the paper to address such concerns. Following Angrist (1995), the problem can be formally stated as follows. The wage-bill is determined through a linear equation conditional on the existence of a private school

$$WB_i = \alpha + \beta GSS_i + \varepsilon_i \tag{6}$$

and a censoring equation (denoting $WB_i = I$ as the indicator for whether WB_i is non-missing)

$$WB_i = I\{\delta GSS_i - \nu_i > 0\}. \tag{7}$$

The instrument, Z_i, determines a first stage

$$GSS_i = \gamma + \mu Z_i + \tau_i. \tag{8}$$

Given the validity of the instrument, Z_i, we assume that $cov(\tau_i, Z_i) = 0$. Then,

$$E(\varepsilon_i|Z_i, WB_i = 1) = E(\varepsilon_i|Z_i, (\delta\gamma + \delta\mu Z_i > \nu_i - \delta\tau_i)$$

so that $cov(\varepsilon_i, Z_i) \neq 0$ in Equation (6) above. Thus, although Z_i is a valid instrument for the decision to setup a private school, it is not a valid instrument in Equation (6). There are two potential solutions.

Following Heckman (1979), if we assume that $(\varepsilon_i, \nu_i, \tau_i)$ are jointly normally distributed, homoskedastic, and independent of Z_i, we obtain the familiar "Mills ratio" as the relevant expectation function conditional on participation. That is,

$$E(\varepsilon_i|Z_i, (\delta\gamma + \delta\mu Z_i > \nu_i - \delta\tau_i) = \lambda(\delta\gamma + \delta\mu Z_i)$$

where $\lambda(\delta\gamma + \delta\mu Z_i) = \frac{-\phi(\lambda(\delta\gamma+\delta\mu Z_i))}{\Phi(\lambda(\delta\gamma+\delta\mu Z_i))}$ and $\phi(.)$ and $\Phi(.)$ are the density and distribution functions of the normal distribution for $\nu_i - \delta\tau_i$. This Mills ratio can is then directly included in Equation (6) as the appropriate selection-correction.

An alternative approach, proposed by Heckman and Robb (1986) and developed by Ahn and Powell (1993), uses the "control-function" approach, where we condition on the predicted probability of $WB_i = 1$ in Equation (6). In essence, this method proposes to estimate β by using pair-wise differences in WB_i for two villages (in our case) for which the non-parametric probability of participation is very close. The approach is implemented by first estimating Equation (7) directly, and then including the predicted probability of participation (and its polynomials) as additional controls in Equation (6).

Appendix II

Comparing Linear IV and Biprobit estimates

Chiburish et al. (2006) show that in the model given by

$$\begin{aligned}
T^* &= \alpha z + c_1 + \varepsilon_1 \\
T &= \mathbf{1}[T \geq 0] \\
Y^* &= \gamma T + c_2 + \varepsilon_2 \\
Y &= \mathbf{1}[Y^* \geq 0]
\end{aligned}$$

with $(\varepsilon_1, \varepsilon_2)$ jointly distributed as standard bivariate normal with correlation ρ, $p_T = (T = 1)$ and $p_Y = (Y = 1)$, the the local average treatment effect or LATE estimated by the linear IV is approximated by

$$\Delta_{LATE} \approx \frac{\gamma}{\sqrt{1-\rho^2}} \phi \left(\frac{\Phi^{-1}(p_Y) - \rho \Phi^{-1}(p_T)}{\sqrt{1-\rho^2}} \right).$$

and the asymptotic variance is approximated by

$$N \, \mathrm{Var}[\hat{\Delta}_{IV}] \approx \frac{p_Y(1-p_Y)}{\alpha^2 [\phi(\Phi^{-1}(p_T))]^2 \mathrm{Var}[z]}.$$

Asymptotic variance of the IV estimator increases as p_Y gets closer to $1/2$ and as p_T gets closer to 0, both of which characterize the case discussed here.

References

[1] Ahn, Hyungtaik, and James Powell. 1993. "Semiparametric Estimation of Censored Selection Models with a Nonparametric Selection Mechanism". *Journal of Econometrics.* 58 (1-2): 3–29.

[2] Alderman, Harold, Jere R. Behrman, Shahrukh Khan, David R. Ross, and Richard Sabot. 1996. "Public Schooling Expenditures in Rural Pakistan: Efficient Targeting Girls and a Lagging Region". In *Public Spending and the Poor: Theory and Evidence*, (ed.) Dominique Van de Walle and Kimberly Nead. Baltimore: Johns Hopkins University Press.

[3] Alderman, Harold, Peter F. Orazem, and Elizabeth M. Paterno. 2001. "School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan". *Journal of Human Resources.* 36 (2): 304–326.

[4] Alderman, Harold, Jooseop Kim, and Peter F. Orazem. 2003. "Design, Evaluation, and Sustainability of Private Schools for the Poor: The Pakistan Urban and Rural Fellowship School Experiments". *Economics of Education Review.* 22 (3): 265–274.

[5] Altonji, Joseph G., Todd E. Elder, and Christoper R. Taber 2005. "An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schools". *Journal of Human Resources*. 40 (4): 791–821.

[6] Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, and Tristan Zajonc. 2006. "Religious Education in Pakistan: A Look at the Data". *Comparative Education Review*. 50 (3): 446.

[7] Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja. 2008. "A Dime a Day: The Possibilities and Limits of Private Schooling in Pakistan". *Comparative Education Review*. 52 (3): 329–355.

[8] Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, Tara Viswanath, and Tristan Zajonc. 2009. *Learning and Educational Achievements in Punjab Schools (LEAPS): Insights to inform the education policy debate*. Forthcoming, Oxford University Press.

[9] Andrabi, Tahir, Jishnu Das, Asim Ijaz Khwaja, and Tristan Zajonc. Forthcoming."Do Value Added Estimates Add Value? Accounting for Learning Dynamics." American Economic Journal: Applied Economics

[10] Andrabi, Tahir, Natalie Bau, Jishnu Das, and Asim Ijaz Khwaja. 2011. "Are bad public schools public bads? Test Scores and Civic Values in private and public schools." In Progress.

[11] Angrist, Joshua David. 1995. "Conditioning on the Probability of Selection to Control Selection Bias". Technical Working Paper No. 181, National Bureau of Economic Research, Cambridge, MA.

[12] Angrist, Joshua, and Victor Lavy. 1999. "Using Maimonedes' Rule to Estimate the Effect of Class Size on Scholastic Achievement". *Quarterly Journal of Economics*. 114 (2): 533–575.

[13] Banerjee, Abhijit V. 2004. "Educational Policy and the Economics of the Family". *Journal of Development Economics*. 74 (1): 3–32.

[14] Banerjee, Abhijit V., and Esther Duflo. 2004. "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program". BREAD Working Paper No. 005, 2004.

[15] Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India". *Quarterly Journal of Economics*. 122 (3): 1235–1264.

[16] Campbell, D. T. 1969. "Reforms as Experiments". *American Psychologist*. 24 (1): 409–429.

[17] Carneiro, Pedro, Jishnu Das, and Hugo Reis. 2010. "Estimating The Demand for School Attributes in Pakistan". In Progress.

[18] Case, Anne, and Angus Deaton. 1999. "School Inputs and Educational Outcomes in South Africa". *Quarterly Journal of Economics*. 114 (3): 1047–1084.

[19] Chaudhury, Nazmul, and Dilip Parjuli. 2010. "Conditional cash transfers and female schooling: the impact of the female school stipend programme on public school enrolments in Punjab, Pakistan". Applied Economics, 1466-4283, Volume 42, Issue 28: 3565 – 3583.

[20] Chaudhury, Nazmul, Jeff Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries". *Journal of Economic Perspectives*. 20 (1): 91–116.

[21] Chiburish, Richard, Jishnu Das, and Misha Lokshin. 2006. "Comparing Linear IV and Bivariate-Probit Estimates". Working Paper, The World Bank.

[22] Corcoran, Sean P., William N. Evans, and Robert M. Schwab. 2004. "Changing Labor Market Opportunities for Women and the Quality of Teachers, 1957-2000". *American Economic Review*. 94 (2): 230–235.

[23] de Janvry, Alain, and Elisabeth Sadoulet. 2006. "Making Conditional Cash Transfer Programs More Efficient: Designing for Maximum Effect of the Conditionality". *World Bank Economic Review*. 20 (1): 1–29.

[24] Downes, Thomas A., and Shane M. Greenstein. 1996. "Understanding the Supply Decisions of Nonprofits: Modeling the Location of Private Schools". *Rand Journal of Economics*. 27 (2): 365–390.

[25] Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment". *American Economic Review*. 91 (4): 795–813.

[26] Duflo, Esther, Rema Hanna, and Stephen Ryan. 2010. "Incentives Work: Getting Teachers to Come to School". Working Paper.

[27] Filmer, Deon, and Norbert Schady. 2008. "Getting Girls into School: Evidence from a Scholarship Program in Cambodia". *Economic Development and Cultural Change*. 56 (3): 581.

[28] Grindle, Merilee Serrill. 2004. *Despite the Odds: The Contentious Politics of Education Reform*. Princeton, NJ: Princeton University Press.

[29] Hamid, Shahnaz. 2010. "Rural to Urban Migration in Pakistan: The Gender Perspective". PIDE Working paper 2010:56.

[30] Heckman, James J. 1979. "Sample Selection Bias as a Specification Error". *Econometrica*, 47 (1): 153–161.

[31] Heckman, James J., and Richard Robb. 1986. "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes". In *Drawing Inferences from Self-Selected Samples*, ed. Howard Wainer. New York: Springer-Verlag.

[32] Hoxby, Caroline M., and Andrew Leigh. 2004. "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States". *American Economic Review*. 94 (2): 236–240.

[33] Jimenez, Emmanuel, Marlaine E. Lockheed, and Vicente Paqueo. 1991. "The Relative Efficiency of Private and Public Schools in Developing Countries". *World Bank Research Observer*. 6 (2): 205–218.

[34] Kim, Jooseop, Harold Alderman, and Peter F. Orazem. 1999. "Can Private School Subsidies Increase Schooling for the Poor? The Quetta Urban Fellowship Program". *World Bank Economic Review*. 13 (3): 443–465.

[35] Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions". *Quarterly Journal of Economics*. 114 (2): 497-532.

[36] Ljungqvist, Lars. 1993. "Economic Underdevelopment: The Case of a Missing Market for Human Capital". *Journal of Development Economics*. 40 (2): 219–239.

[37] Muralidharan, Karthik, and Venkatesh Sundararaman. Forthcoming."Teacher Performance Pay: Experimental Evidence from India". Journal of Political Economy .

[38] Pitt, Mark M., Mark R. Rosenzweig, and Donna M. Gibbons. 1995. "The Determinants and Consequences of the Placement of Government Programs in Indonesia". In *Public Spending and the Poor: Theory and Evidence*, (ed.) Dominique Van de Walle and Kimberly Nead. Baltimore: Johns Hopkins University Press.

[39] Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. “Teachers, Schools, and Academic Achievement”. *Econometrica*. 73 (2): 417–458.

[40] Schultz, T. Paul. 2004. “School Subsidies for the Poor: Evaluating the Mexican Progresa poverty program”. *Journal of Development Economics*. 74 (1): 199–250.

[41] Stock, James H., Jonathon H. Wright, and Motohiro Yogo. 2002. “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments”. *Journal of Business and Economic Statistics*. 20 (4): 518–529.

[42] Tilak, Jandhyala B. G., and Ratna M. Sudarshan. 2002. “Private Schooling in Rural India”. Working Paper #WP010009, National Council of Applied Economic Research, New Delhi, India.

[43] Tooley, James. 2005. “Private Schools for the Poor: Education Where No One Expects It”. *Education Next*. 5 (4): 22–32.

[44] Tooley, James, and Pauline Dixon. 2005. *Private Education is Good for the Poor: A study of private schools serving the poor in low-income countries*. Washington, DC: CATO Institute.

[45] UNESCO. 2004. *Initiative for Teacher Education in Sub-Saharan Africa*. New York: United Nations Educational, Scientific and Cultural Organization.

[46] Urquiola, Miguel. 2006. “Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia”. *The Review of Economics and Statistics*. 88 (1): 171–177.

[47] Urquiola, Miguel, and Emiliana Vegas. 2005. “Arbitrary Variation in Teacher Salaries: An Analysis of Teacher Pay in Bolivia.” In *Incentives to Improve Teaching*, (ed.) Emiliana Vegas. Washington, DC: The World Bank.

[48] World Bank. 2005. *Improving Gender Outcomes: The Promise for Pakistan*. Washington, DC: The World Bank.

Table I - Summary Statistics (Village Level)

Variable	*Mean*	*50th Percentile*	*S.D.*	*N*
1981 Number of Women with Middle+ Education	4.28	1	17.94	9333
1998 Number of Women with Middle+ Education	26.74	9	92.80	9333
1981 Percent Women with Middle+ Education	0.01	0	0.03	8882
1998 Percent Women with Middle+ Education	0.06	0.03	0.07	8915
Households Per Capita With Radio Access (1998)	0.03	0.02	0.03	8952
Ratio of Females to Males, Under Age 14 (1998)	0.94	0.93	0.24	8892
Area (Acres, 1998)	1550.34	1042	2520.51	9091
Percent of Houses Permanent (1998)	0.06	0.06	0.05	8935
1981 Total Population	1020.36	667.00	1247.91	9333
1998 Total Population	1537.70	961.00	2053.87	9333
1981 Population of Largest Village in PC	1670.04	1375.00	1310.46	9333
Number of Villages in PC (1998)	4.57	4	2.28	9333
Girls' Secondary School Exists	0.02	0	0.16	9333
Girls' Primary School Exists	0.56	1	0.50	9330
Boys' Secondary School Exists	0.01	0	0.12	9333
Boys' Primary School Exists	0.70	1	0.46	9330
Girls' Secondary School Exposure (if one exists)	13.15	14	5.47	232
Girls' Primary School Exposure (if one exists)	21.43	18	11.80	4967
Boys' Secondary School Exposure (if one exists)	12.62	13.50	5.16	138
Boys' Primary School Exposure (if one exists)	35.21	31	19.66	6475
Private School Exists	0.13	0	0.33	9258
Number of Private Schools	0.22	0	0.87	9258
Private School Enrollment Rate (if one exists)	0.12	0.06	0.37	1165

This table presents summary statistics for various variables of interest. The years for which the above data are given varies by source: All 1981/1998 variables are from the 1981/1998 Population Censuses while all schooling data is from the EMIS, NEC, or Private School Census.

Table II - First Stage and Falsification Tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	First Stage - Probit and OLS		OLS Falsification Tests - Other Public Goods						Falsification Test - Probit with QH Top Rank	
Dependent Variable	Girls' Secondary School	Girls' Secondary School	Girls' Primary School	Boys' Primary School	Boys' Secondary School	Water	Electricity	Permanent Houses	Girls' Secondary School	Girls' Secondary School
Instrument	0.0124***	0.016***	0.011	0.011	0.0008	-0.0002	0.002	0.002		0.012***
	(0.004)	(0.005)	(0.015)	(0.014)	(0.0037)	(0.0008)	(0.001)	(0.001)		(0.004)
Has Highest Population									-0.005	-0.003
in QH, 1981									-0.008	(0.008)
Girls' Primary School Exists		-0.052***								
		(0.004)								
Boys' Secondary School Exists		0.232***								
		(0.017)								
Boys' Primary School Exists		0.003								
		(0.005)								
Area (000s of Acres)		0.001								
		(0.002)								
% Houses Permanent		0.076								
		(0.053)								
1981 Population (000s)	0.0059*	0.014	0.293***	0.374***	-0.002	0.002	0.018***	-0.002	0.0087**	0.006*
	(0.004)	(0.010)	(0.027)	(0.025)	-0.007	(0.001)	(0.002)	-0.002	(0.0035)	(0.004)
1981 Population (000s)^2	-0.0003	(0.001)	-0.033***	-0.041***	0.002*	-0.0003*	-0.002***	0.0001	-0.0005**	-0.0003
	(0.0002)	(0.001)	(0.003)	(0.003)	(0.001)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
1998 Population (000s)	0.003	0.003	-0.019	-0.054***	0.008**	0.0001	-0.002	0.001	0.0029	0.003
	(0.002)	(0.005)	(0.014)	(0.013)	(0.003)	(0.0007)	(0.001)	-0.001	(0.0020)	(0.002)
1998 Population (000s)^2	$-2\text{x}10^{-5}$	0.0002	0.002***	0.003***	$3\text{x}10^{-5}$	$2\text{x}10^{-5}$	$1\text{x}10^{-4}$***	$-1\text{x}10^{-5}$	$-2\text{x}10^{-5}$	$-2\text{x}10^{-5}$
	($5\text{x}10^{-5}$)	(0.0001)	(0.0004)	(0.0004)	($10\text{x}10^{-5}$)	($2\text{x}10^{-5}$)	($0.1\text{x}10^{-4}$)	($3\text{x}10^{-5}$)	($5\text{x}10^{-5}$)	($5\text{x}10^{-5}$)
PC Fixed Effects		Y	Y	Y	Y	Y	Y	Y		
R-Squared			0.48	0.46	0.48	0.42	0.70	0.69		
Adjusted R-Squared		0.17								
Pseudo R-Squared	0.05								0.05	0.05
Chi-Square Stat (Instrument =	13.04									
F-Stat (Instrument = 0)		8.9								
Observations	9333	8705	9330	9330	9333	8935	8935	8935	9333	9333

Standard errors in parentheses with * indicating significance at 10%, ** at 5%, and *** at 1%. Columns (1)-(2) present first stage regressions using the eligibility rule as a predictor for the location of GSS. Column (1) gives the increased probability of finding a GSS in an eligible village (with basic population controls). Column (2) presents a linear first stage that includes controls for the village's population in 1981 and 1998, other village level public goods, and PC fixed effects. Columns (3)-(8) check that the instrument does not predict other public goods. Columns (9)-(10) show that a village having the highest population within a QH does not predict GSS construction.

Table III - GSS Impact on Private School Existence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	First Difference	Linear, Second Stage	Bivariate Probit - ATE	Bivariate Probit - Average ToT Effect	Reduced Form - Program QHs	Reduced Form - Non-Program QHs
Instrument						0.038***	-0.0017
						(0.014)	(0.016)
Girls' Secondary School Exists	0.095***	0.097***	1.505*	0.266*	0.246***		
(= Received GSS After 1981)	(0.025)	(0.025)	(0.802)	(0.151)	(0.092)		
Girls' Primary School Exists	0.016*		0.089**	xx	xx	0.007	0.017
	(0.0080)		(0.043)			(0.011)	(0.014)
Boys' Secondary School Exists	-0.005		-0.333*	xx	xx	-0.030	0.112*
	(0.034)		(0.191)			(0.040)	(0.063)
Boys' Primary School Exists	-0.005		-0.009	xx	xx	-0.009	0.001
	(0.009)		(0.012)			(0.012)	(0.014)
Received Girls' Primary		0.0190					
School After 1981		(0.035)					
Received Boys' Secondary		-0.011					
School After 1981		(0.008)					
Received Boys' Primary		-0.026***					
School After 1981		(0.010)					
Area (000s of Acres)	-0.008**		-0.009**	xx	xx	-0.029***	-0.002
	(0.003)		(0.004)			(0.007)	(0.004)
% Houses Permanent	0.194*		0.083	xx	xx	0.184	0.208
	(0.103)		(0.142)			(0.133)	(0.163)
1981 Population (000s)	0.046***		0.013	xx	xx	0.035	0.054*
	(0.017)		(0.028)			(0.027)	(0.028)
1981 Population $(000s)^2$	-0.0030		-0.0002	xx	xx	0.004	-0.007**
	(0.002)		(0.003)			(0.004)	(0.003)
1998 Population (000s)	0.064***		0.059***	xx	xx	0.060***	0.067***
	(0.009)		(0.012)			(0.012)	(0.014)
1998 Population $(000s)^2$	-0.001***		-0.001***	xx	xx	-0.002***	-0.002***
	(0.0003)		(0.0004)			(0.0005)	(0.0004)
Δ Population (000s)		0.075***					
		(0.005)					
PC Fixed Effects	Y	Y	Y			Y	Y
R-Squared						0.51	0.57
Adjusted R-Squared	0.31	0.28					
Prob > F						37.81	24.71
Prob > Chi-Square			0.00				
Observations	8705	8900	8705	8705	8705	5191	3514
Number of PCs (1998)			2784				

Standard errors in parentheses with * indicating significance at 10%, ** at 5%, and *** at 1%. This table presents regression results for which the dependent variable is a dummy indicating the presence of at least one private school in a village (or the change in this variable for the first difference specification). Column (1) gives OLS results for the impact of GSS on private school existance. Column (2) shows a first-differenced specification. (First-differencing Girls' Seconary School Exists does not change the variable because our sample contains no villages which had a GSS prior to 1981. That is, having a GSS in our sample is equivalent to receiving one after 1981.) Columns (3)-(5) present the IV specifications. Column (3) gives the second stage results from a linear specification. Columns (4)-(5) implement the bivariate probit specification and report, respectively, the average treatment effect and the treatment on the treated effect of a GSS on the existence of a private school with analytical standard-errors computed using the delta method. Controls are present in these two regressions where marked, but coefficients and standard errors are not given. These regressions also include (in the absence of PC fixed effects) linear and quadratic controls for the population of the largest village in the PC as well as a control for the number of villages in the PC. Columns (6)-(7) present an additional check of the instrument by showing that the reduced form only holds in broad areas where at least one GSS was provided. Villages are divided into two sub-groups: "program regions," where at least one village in the QH received a GSS [Column(6)]; and "non-program regions," where no village in the QH received a GSS [Column (7)].

Table IV - Private School Existence: The Female Teacher Channel

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	Private School Exists		Number of Women with Middle+ Education				Private School Exists	
	OLS	First Difference	OLS	First Difference	Second Stage	Reduced Form	OLS	First Difference
Instrument						3.46***		
						(1.19)		
Years Exposure to Girls' Secondary School	0.006***							
	(0.002)							
Years Exposure to Girls' Primary School	0.0015***							
	(0.000)							
Years Exposure to Boys' Secondary School	0.001							
	(0.003)							
Years Exposure to Boys' Primary School	0.0004**							
	(0.0002)							
Girls' Secondary School Exists		0.097***	10.81***	9.52***	219.32**			
		(0.025)	(2.93)	(3.55)	(103.06)			
Girls' Primary School Exists			2.37**		13.08**	1.79*		
			(0.99)		(5.46)	(0.98)		
Boys' Secondary School Exists			7.51*		-40.96*	9.98**		
			(3.98)		(24.55)	(3.93)		
Boys' Primary School Exists			1.28		0.59	1.27		
			(1.06)		(1.49)	(1.06)		
Received Girls' Primary School After 1981		-0.011		-4.32***				
		(0.008)		(1.17)				
Received Boys' Secondary School After 1981		0.019		14.25***				
		(0.035)		(4.91)				
Received Boys' Primary School After 1981		-0.026***		-0.65				
		(0.010)		(1.36)				
% Women with Middle+ Education							0.376***	
							(0.084)	
% Men with Middle+ Education							0.033	
							(0.049)	
Δ % Women w/ Middle+ Education								0.414***
								(0.086)
Δ % Men w/ Middle+ Education								-0.047
								(0.050)
Area (000s of Acres)	-0.008**		-2.03***		-2.15***	-2.03***	-0.008**	
	(0.003)		(0.39)		(0.53)	(0.39)	(0.003)	
% Houses Permanent	0.187*		44.83***		28.43	45.14***	0.276**	
	(0.104)		(12.05)		(18.30)	(12.06)	(0.128)	
1981 Population (000s)	0.028		-1.61		-6.36*	-3.35	0.046***	
	(0.018)		(2.04)		(3.63)	(2.15)	(0.017)	
1981 Population (000s)^2	-0.001		-0.32		0.07	-0.16	-0.003	
	(0.002)		(0.24)		(0.38)	(0.25)	(0.002)	
1998 Population (000s)	0.065***		9.71***		8.91***	9.45***	0.064***	
	(0.009)		(1.05)		(1.49)	(1.06)	(0.009)	
1998 Population (000s)^2	-0.0012***		1.79***		1.76***	1.79***	-0.0012***	
	(0.0003)		(0.03)		(0.05)	(0.03)	(0.0003)	
Δ Population (000s)		0.075***		60.39***				0.072***
		(0.005)		(0.71)				(0.005)
PC Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Adjusted R-Squared	0.32	0.28	0.88	0.77		0.88	0.32	0.27
Prob > F					4.08			
Observations	8355	8900	8705	8975	8705	8705	8685	8711

Standard errors in parentheses with * indicating significance at 10%, ** at 5%, and *** at 1%. Columns (1)-(2) present estimates for the effects of school exposure on private school existance from a linear probability model and a first difference specification. (First-differencing Girls' Seconary School Exists does not change the variable because our sample contains no villages which had a GSS prior to 1981. That is, having a GSS in our sample is equivalent to receiving one after 1981.) Using the same approach, Columns (3)-(4) assess the correlation between educated women and the presence of a GSS. Columns (5)-(6) examine this relationship through an instrumental variable specification and present the second stage and reduced form. Finally, columns (7)-(8) show the extent to which the extent of secondary-school-educated women in the village are associated with private school existence.

Table V - Supply Side Impact: Teaching Costs

	(1)	(2)	(3)	(4)	(5)	(7)	(6)
	OLS	Heckman	Heckman (Expanded First Stage)	Control Function	Control Function (Expanded First Stage)	OLS	OLS
Girls' Secondary School Exists	-0.318*	-0.321***	-0.321***	-0.324*	-0.325*	-6.488*	-0.614**
	(0.186)	(0.091)	(0.092)	(0.191)	(0.187)	(3.370)	(0.309)
Girls' Primary School Exists	0.075	0.061	0.061	0.069	0.068	0.057	0.074
	(0.087)	(0.042)	(0.043)	(0.099)	(0.088)	(0.087)	(0.087)
Boys' Secondary School Exists	0.295	0.282**	0.285**	0.269	0.269	0.333	0.299
	(0.220)	(0.111)	(0.112)	(0.226)	(0.225)	(0.221)	(0.220)
Boys' Primary School Exists	0.019	0.015	0.0001	0.013	0.010	0.036	0.011
	(0.087)	(0.042)	(0.044)	(0.096)	(0.091)	(0.087)	(0.087)
Ratio of Females to Males, Under Age 14						0.118	
						(0.353)	
Ratio of Females to Males, Under Age 14 × Girls' Secondary School Exists						6.492*	
						(3.536)	
Households Per Capita With Radio Access							1.541
							(1.414)
Households Per Capita With Radio Access × Girls' Secondary School Exists							9.223
							(8.161)
Area (000s of Acres)	-0.058	-0.056***	-0.058***	-0.058	-0.058	-0.061	-0.061
	(0.039)	(0.021)	(0.021)	(0.040)	(0.039)	(0.039)	(0.039)
% Houses Permanent	0.006	0.055	0.046	0.016	0.003	-0.058	-0.238
	(1.320)	(0.635)	(0.644)	(1.329)	(1.327)	(1.318)	(1.333)
1981 Population (000s)	0.122	0.274**	0.253**	0.198	0.186	0.138	0.136
	(0.104)	(0.110)	(0.105)	(0.185)	(0.182)	(0.105)	(0.105)
1981 Population $(000s)^2$	-0.021*	-0.039***	-0.037***	(0.031)	(0.030)	-0.021*	-0.022*
	(0.012)	(0.014)	(0.013)	(0.021)	(0.021)	(0.012)	(0.012)
1998 Population (000s)	0.028	0.097*	0.088*	0.033	0.027	0.016	0.027
	(0.053)	(0.052)	(0.049)	(0.093)	(0.092)	(0.053)	(0.053)
1998 Population $(000s)^2$	0.001	-0.001	-0.001	0.0002	0.0003	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
PC Fixed Effects	Y	Y	Y	Y	Y	Y	Y
Adjusted R-Squared	0.46			0.45	0.46	0.46	0.46
Observations	1090	9292	9292	1090	1090	1090	1090

Standard errors in parentheses with * indicating significance at 10%, ** at 5%, and *** at 1%. This table examines the impact of GSS on skilled women wages. The dependent variable is the (logarithm of the) average salary of a private school teacher in the village. Since private school teachers are almost entirely women and educated women are mostly employed as teachers, this measure is a reasonable proxy for skilled women wages. Column (1) presents the OLS results. The sample is slightly smaller than the number of villages where there is a private school since, in a few cases in the PEIP data, private schools did not report wages. Columns (2)-(5) correct for selection into the wage sample. Columns (2)-(3) present results using Heckman's selection model. Columns (4)-(5) use the "control function" approach. Columns (3) and (5) include the presence of a government boys primary school in the village as an additional instrument in the selection stage. Finally, columns (6)-(7) present tentative evidence that wage declines are larger in villages where labor markets for women are more restricted. Column (6) examines the differential effect of GSS construction on wages for more and less progressive villages using the female/male ratio for children under the age of 14 as an indicator of gender bias. Column (7) presents similar results using households per capita with access to radios as an indicator of village-level development.

Figure I - Private School Existence / Rule-Based Instrument and 1981 Population

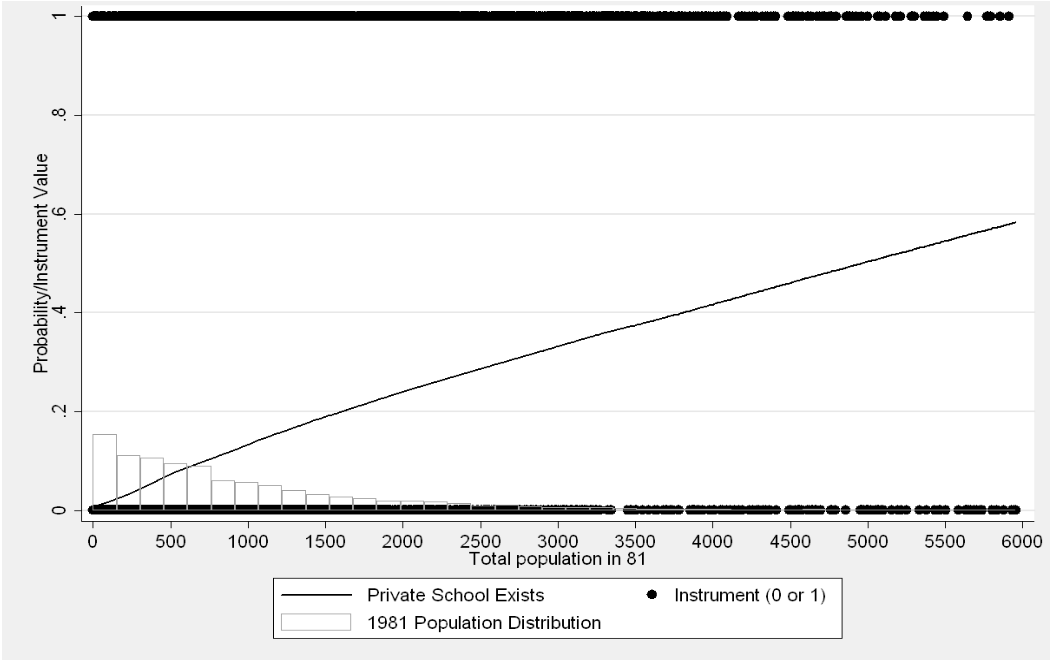

Figure I illustrates how the existence of private schools and the binary instrument covary with 1981 village population (the relationship with 1998 population is very similar). Here, we plot the binary instrument, $Rule_i$, for all villages in our sample and the non-parametric relationship between private school location and village population. We note that there are both "eligible" and "ineligible" villages at all population levels. The bar graphs illustrates the population distribution.

Figure II - Probabilities of Schools Existing by Instrument and Population Quartiles

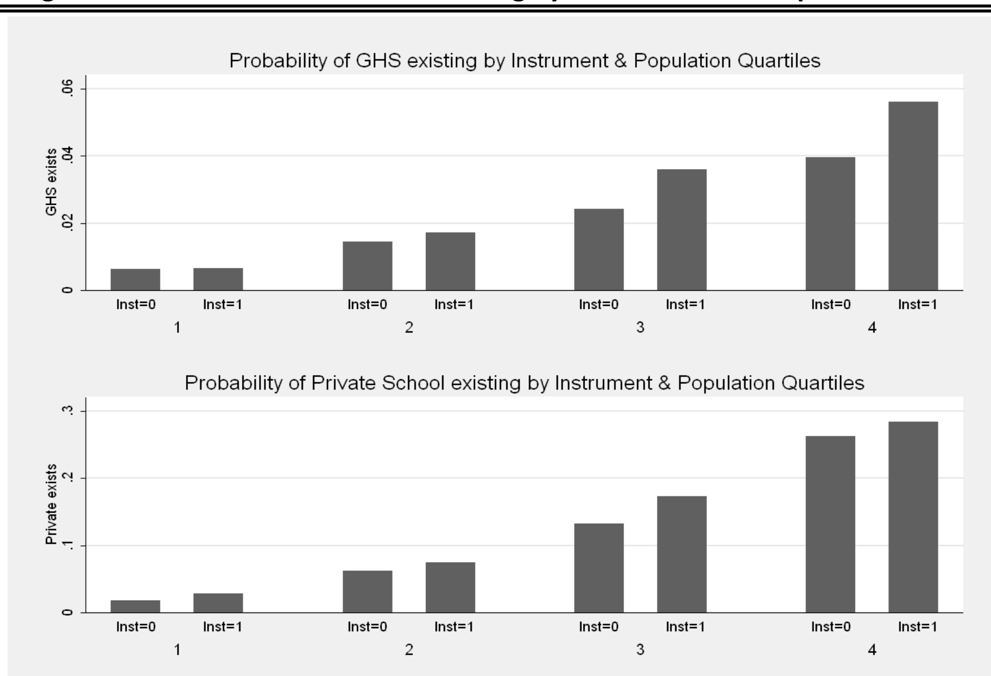

Figure II provides a simple illustration of the our instrumental variable estimates by dividing villages into four population quartiles, averaged over 1981 and 1998 populations. The top panel illustrates the first stage by comparing the percentage of villages with a GSS in the "eligible" group compared to the "ineligible" group. The bottom panel illustrates the reduced form, by comparing, over the same population quartiles, the percentage of villages with a private school in the "eligible" and "ineligible" groups.

Appendix Table I - Differences in Means

Variable	*Instrument=1*	*Instrument=0*	*Difference*	*P-Value*
Area in Acres (1998)	2084.61	1326.88	757.73	0.00
	44.58	32.12	57.43	
Total Population (1981)	1644.75	759.79	884.96	0.00
	22.78	14.48	26.82	
Total Population (1998)	2516.91	1129.06	1387.85	0.00
	43.42	22.22	44.38	
% Δ Population (1981 to 1998)	0.62	0.69	-0.07	0.24
	0.030	0.037	0.059	
Ratio of Females to Males (1981)	0.904	0.904	0.000	0.99
	0.006	0.004	0.007	
Ratio of Females to Males (1998)	0.938	0.946	-0.007	0.16
	0.005	0.003	0.005	
% Women Aged 4 and Below (1981)	0.158	0.154	0.004	0.63
	0.007	0.005	0.008	
% Women Aged 5-14 (1981)	0.285	0.284	0.001	0.92
	0.009	0.006	0.010	
% Women with ID Card (1998)	0.490	0.478	0.012	0.30
	0.010	0.006	0.012	
% Literate Women, Aged 15+ (1981)	0.016	0.017	-0.001	0.74
	0.002	0.002	0.003	
% Women with Middle+ Education, Aged 15+	0.014	0.014	0.000	0.91
	0.002	0.001	0.003	
% Men Aged 4 and Below (1981)	0.144	0.141	0.004	0.65
	0.007	0.004	0.008	
% Men Aged 5-14 (1981)	0.293	0.291	0.003	0.81
	0.009	0.006	0.010	
% Men with ID Card (1998)	0.691	0.684	0.007	0.50
	0.009	0.006	0.011	
% Literate Men, Aged 15+ (1981)	0.169	0.166	0.003	0.73
	0.007	0.005	0.009	
% Men with Middle+ Education, Aged 15+ (1981)	0.120	0.119	0.001	0.95
	0.006	0.004	0.007	
% Houses Permanent (1998)	0.063	0.065	-0.002	0.76
	0.005	0.003	0.006	
% Households with Water (1998)	0.011	0.010	0.001	0.61
	0.002	0.001	0.002	
% Households with Electricity (1998)	0.075	0.068	0.006	0.27
	0.005	0.003	0.006	
% Households with TV (1998)	0.029	0.028	0.001	0.82
	0.003	0.002	0.004	
% Household with Radio (1998)	0.025	0.028	-0.003	0.38
	0.003	0.002	0.004	

Standard errors in parentheses. This tables gives evidence that there are no unexpected baseline differences in observables between eligible (Instrument = 1) and ineligible (Instrument = 0) villages. The only significant differences are in population and area, which arise directly from the construction of the instrument. Several 1998 variables of interest are included when 1981 numbers are not available, though these are not, strictly speaking, baseline measurements.

Appendix Table II - Full Sample Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	First Stage / Falsification Test (OLS)		Impact on Private School Existence			Channels (OLS)			Wages
	Girls' Secondary School	Boys' Secondary School	OLS	Linear, Second Stage	Bivariate Probit - ATE	Private School Existence	Number of Women with Middle+ Education	Private School Existence	Heckman
Instrument	0.037***	0.058***							
	(0.006)	(0.007)							
Girls' Secondary School Exists			0.100***	1.082***	0.309***		31.82***		
			(0.009)	(0.257)	(0.033)		(1.41)		
Girls' Primary School Exists	-0.227***		-0.007	0.217***	xx		-3.19***		
	(0.005)		(0.006)	(0.059)			(0.94)		
Boys' Secondary School Exists	0.254***		0.093***	-0.158**	xx		8.36***		
	(0.007)		(0.008)	(0.067)			(1.35)		
Boys' Primary School Exists	0.043***		-0.003	-0.045***	xx		0.35		
	(0.006)		(0.006)	(0.014)			(1.00)		
Years Exposure to Girls' Secondary School						0.003***			-0.003*
						(0.0002)			(0.001)
Years Exposure to Girls' Secondary School^2									0.0001**
									(0.000)
Years Exposure to Girls' Primary School						0.001***			-0.006***
						(0.0002)			(0.002)
Years Exposure to Girls' Primary School^2									0.0001***
									(0.000)
Years Exposure to Boys' Secondary School						0.0016***			-0.0001
						(0.0001)			(0.001)
Years Exposure to Boys' Secondary School^2									-1x10^{-9}
									(1x10^{-9})
Years Exposure to Boys' Primary School						0.0004***			-0.0006
						(0.0001)			(0.001)
Years Exposure to Boys' Primary School^2									1x10^{-9}
									(1x10^{-9})
% Women with Middle+ Education								0.589***	
								(0.055)	
% Men with Middle+ Education								0.090**	
								(0.035)	
Area (000s of Acres)	-0.005***		-0.009***	-0.004*	xx	-0.009***	-3.85***	-0.008***	-0.000***
	(0.002)		(0.002)	(0.003)		(0.002)	(0.28)	(0.002)	(0.000)
% Houses Permanent	0.123**		0.331***	0.209**	xx	0.332***	38.82***	0.349***	0.425
	(0.059)		(0.067)	(0.094)		(0.067)	(10.78)	(0.088)	(0.325)
1981 Population (000s)	0.037***	0.054***	0.059***	-0.004	xx	0.053***	-0.37	0.068***	0.040
	(0.004)	(0.004)	(0.007)	(0.019)		(0.008)	(1.16)	(0.007)	(0.029)
1981 Population (000s)^2	-0.0006***	-0.001***	-0.004***	(0.001)	xx	-0.005***	0.85***	-0.005***	-0.005*
	(0.0001)	(0.0001)	(0.001)	(0.001)		(0.001)	(0.13)	(0.001)	(0.003)
1998 Population (000s)	0.053***	0.081***	0.086***	0.044'***	xx	0.082***	21.34***	0.094***	0.024
	(0.007)	(0.007)	(0.004)	(0.012)		(0.004)	(0.64)	(0.004)	(0.015)
1998 Population (000s)^2	-0.003***	-0.003***	-0.002***	-0.001***	xx	-0.002***	0.99***	-0.002***	1x10^{-9}
	(0.001)	(0.001)	(0.0001)	(0.0003)		(0.0002)	(0.02)	(0.0001)	(0.0003)
PC Fixed Effects	Y	Y	Y	Y		Y	Y	Y	Y
R-Squared		0.48							
Adjusted R-Squared	0.40		0.38			0.38	0.86	0.38	
F-Stat (Instrument = 0)	33.80								
Prob > Chi-Square				0.00					
Observations	23756	25874	23756	23756	23756	22845	23756	23698	27819
Number of PCs (1998)				7142					

Standard errors in parentheses with * indicating significance at 10%, ** at 5%, and *** at 1%. This table replicates some of the main regressions in the previous tables to demonstrate that the results hold in the full sample as well. Column (1) and (2) correspond to Table II, Columns (2) and (5), respectively. Column (3), (4), and (5) correspond to Table III, Columns (1), (3), and (4), respectively. Columns (6), (7), and (8) correspond to Table IV, Columns (1), (3), and (7), respectively. Column (9) corresponds to Table V, Column (2). Column (9) includes squared terms for exposure in the expectation that short-term exposure decreases wages by increasing supply, while in the longer term, exposure may increase wages as educated women become mothers who increase demand for teachers.