School Inputs, Household Substitution, and Test Scores

Jishnu Das Stefan Dercon James Habyarimana Pramila Krishnan
Karthik Muralidharan Venkatesh Sundararaman

1 November 2011*

Abstract: Empirical studies of the relationship between school inputs and test scores typically do not account for the fact that households will respond to changes in school inputs. We present a dynamic household optimization model relating test scores to school and household inputs, and test its predictions in two very different low-income country settings – India and Zambia. We measure student test score gains in response to unanticipated as well as anticipated changes in school funding and find that unanticipated school grants lead to significant improvements in student test scores but anticipated grants have no impact on test scores. We show that the most likely mechanism for this result is that households offset anticipated grants more than unanticipated grants. Our results suggest that (a) household responses will mediate the impact of school spending on learning outcomes, and (b) naïve estimates of public education spending on learning outcomes that do not account for optimal household responses are likely to be biased if used to estimate parameters of an education production function.

JEL Classification: H52, I21, O15

Keywords: school grants, school inputs, household substitution, education in developing countries, randomized experiment, India, Zambia, Africa, education production function

---

* Jishnu Das (World Bank and Center for Policy Research, Delhi: jdas1@worldbank.org)
Stefan Dercon (Oxford University, BREAD, and CEPR: stefan.dercon@economics.ox.ac.uk)
James Habyarimana (Georgetown University, IZA, and Center for Global Development: jph35@georgetown.edu)
Pramila Krishnan (Cambridge University and CEPR: pk237@cam.ac.uk)
Karthik Muralidharan (UC San Diego, NBER, BREAD, and J-PAL: kamurali@ucsd.edu)
Venkatesh Sundararaman (World Bank: vsundararaman@worldbank.org)

We thank Julie Cullen, Gordon Dahl, Roger Gordon, Gordon Hanson, Hanan Jacoby, Andres Santos, and several seminar participants for comments. The World Bank and the UK Department for International Development (DFID) provided financial support for both the Zambia and India components of this paper. The experiment in India is part of a larger project known as the Andhra Pradesh Randomized Evaluation Study (AP RESt), which is a partnership between the Government of Andhra Pradesh, the Azim Premji Foundation, and the World Bank. We thank Dileep Ranjekar, Amit Dar, Samuel C. Carlson, and officials of the Department of School Education in Andhra Pradesh for their continuous support. We are especially grateful to DD Karopady, M Srinivasa Rao, and staff of the Azim Premji Foundation for their leadership in implementing the project in Andhra Pradesh. Vinayak Alladi and Ketki Sheth provided excellent research assistance. The findings, interpretations, and conclusions expressed in this paper are those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the governments they represent.

1. Introduction

The relationship between school inputs and education outcomes is of fundamental importance for education policy and has been the subject of hundreds of empirical studies around the world (see Hanushek 2002, and Hanushek and Luque 2003 for reviews of US and international evidence respectively). However, while the empirical public finance literature has traditionally paid careful attention to the behavioral responses of agents to public programs[1], the empirical literature estimating education production functions has rarely accounted for household re-optimization in response to public spending. This is a critical gap because (a) household responses to education policies will mediate the extent to which different types of education spending translate into learning outcomes, and (b) parameters of education production functions are typically not identified if household inputs respond to changes in school-level inputs (see Urquiola and Verhoogen 2009 for one such example in the context of class-size).

We develop a dynamic model of household optimization that clarifies how increases in school-provided inputs and household spending responses translate into learning outcomes. We then test the main predictions of the model in two very different countries – India and Zambia – using panel data on student achievement combined with unique matched data sets of school and household spending. A key contribution of this paper is our ability to measure changes in household spending as well as student test-score gains in response to both *unanticipated* as well as *anticipated* changes in school funding. The former measures the production function effect of increased school funding (a partial derivative holding other inputs constant), while the latter measures the policy effect (a total derivative that accounts for re-optimization by agents).

The theoretical framework of a dynamic forward-looking model provides a useful guide to the key issues. In this framework, households' optimal spending decisions will take into account all information available at the time of decision making. The impact of school inputs on test scores depends then on (a) whether such inputs are anticipated or not and (b) the extent of substitutability between household and school inputs in the education production function. The model predicts that if household and school inputs are technical substitutes, an anticipated increase in school inputs in the next period will decrease household contributions that period. Unanticipated increases in school inputs limit the scope for household responses, leaving

[1]Illustrative examples include Meyer (1990) on unemployment insurance, Cutler and Gruber (1996) on health insurance, Eissa and Leibman (1996) on the EITC, Autor and Duggan (2003) on disability insurance. See Moffitt (2002) for an overview on labor supply responses to welfare programs.

household contributions unchanged in the short run. These differences lead to a testable prediction: If household and school inputs are (technical) substitutes, unanticipated inputs will have a larger impact on test scores than anticipated inputs. We test this prediction using data on education spending for largely substitutable school inputs, such as books and writing materials in both India and Zambia.

Our main test of the model is based on experimental variation in school funds induced by a randomly-assigned school grant program in the Indian state of Andhra Pradesh (AP). The AP school block grant experiment was conducted across a representative sample of 200 government-run schools in rural AP with 100 schools selected by lottery to receive a school grant (worth around $3 per pupil) over and above their regular allocation of teacher and non-teacher inputs. The conditions of the grant specified that the funds were to be spent on inputs used directly by students and not on any infrastructure or construction projects. The program was implemented for two years. In the first year, the grant (which was assigned by lottery) was a surprise for recipient schools while in the second year, the grant was anticipated by the parents and teachers of program schools.

Our strongest results are that household education spending in program schools does not change in the first year (relative to spending in the control schools), but that it is significantly lower in the second year suggesting that households offset the anticipated grant significantly more than they offset the unanticipated grant. Evaluated at the mean, we find that for each dollar provided to treatment schools in the second year, household spending declines by 0.76 dollars. We cannot reject that the grant is completely offset by the household, while the lower bound of a 95% confidence interval suggests that at least half is crowded out. In short, we find considerable crowding out of the school grant by households in the second year.

Consistent with this, we find that students in program schools perform significantly better than those in comparison schools at the end of the first year of the school grant program, scoring 0.08 and 0.09 standard deviations (SD) more in language and mathematics tests respectively for a transfer of a little under $3 per pupil. In the second year, the treatment effects of the program are considerably lower and not significantly different from zero. These results suggest that the production-function effect of the school grants on test scores was positive, but that the policy effects are likely to be lower once households re-optimize (and converge to the income elasticity of test scores if the materials provided in the school are fully substitutable).

While the experiment in AP provides stronger identification, our data from Zambia allow us to consider the model in the context of a scaled up program implemented across an entire country by a national government. Starting in 2001, the Government of Zambia started providing all schools in the country with a fixed block grant of $600-650 (regardless of enrollment) as part of a nationally well-publicized program. Thus, variation in school enrollment led to substantial cross-sectional variation in the per-student funding provided by this rule-based grant. We find, however, that per-student variation in the block grant is not correlated with any differences in student test score gains. As in AP, we collect data on household spending and find that household spending almost completely offsets variations in predicted per-student school grants, suggesting that household offset may have been an important channel for the lack of correlation between public education spending and test score gains . We further exploit the presence of a discretionary district-level source of funding that is highly variable across schools and much less predictable than the rule-based grant and find that student test scores in schools receiving these funds are 0.10 standard deviations higher for both the English and mathematics tests for a median transfer of just under $3 per pupil.

These findings are consistent with the predictions of the model and are virtually identical to those from AP. The two sets of results complement each other and provide greater external validity to our findings. The AP case offers experimental variation in one source of funding, which changes from being unanticipated to anticipated over time. The Zambia case offers an analysis of two contemporaneously different sources of funding (rule-based and discretionary) in a scaled up government-implemented setting, but relies on non-experimental data.

There are important policy implications of our results. The impact of anticipated school grants in both settings is low, not because the money did not reach the schools (it did) or because it was not spent well (there is no evidence to support this), but because households realigned their own spending patterns optimally across time and other sources of spending, and not just on their children's education. The replication of the findings in two very different settings[2], with

[2] The two settings are similar in some ways including having high primary school enrollment but low student test scores and having limited funding for recurrent non-salary expenditures (Pratham 2010, Kanyika et al. 2005). However, at the time of the study, Zambia experienced severe declines in per-capita government education expenditure and a stagnant labor market, while Andhra Pradesh has been one of the fastest growing states in India with large increases in government spending in education over the last decade. Our finding very similar results in a dynamic, growing economy and in another that was, at best, stagnant at the time of our study suggests that the results generalize across very different labor market conditions and the priority given to education in the government's budgetary framework.

two different implementing agencies (a leading non-profit organization in AP, and the government in Zambia), and in representative population-based samples suggests that the impact of school grant programs is likely to be highly attenuated by household responses. This has direct implications for thinking about the effectiveness of many such school grant programs across several developing countries.[3]

The distinction between anticipated and unanticipated inputs and the differential ability of households to substitute across various inputs may account for the wide variation in estimated coefficients of school inputs on test scores (Glewwe 2002, Hanushek 2003, or Kreuger 2003), and our results highlight the empirical importance of distinguishing between policy effects and production function parameters (Todd and Wolpin 2003, and Glewwe and Kremer 2005 make this point theoretically). A failure to reject the null hypothesis in studies that use the production function approach could arise either because the effect of school inputs on test scores through the production function is zero or because households (or teachers or schools) substitute their own resources for such inputs. While in our case the substitution takes the form of textbooks or writing materials, in a more general setting it may include parental time[4], private tuition and other inputs.[5] Our results show that the policy effect of school inputs is different from the production function parameters with consequences both for estimation techniques and for policy (see Pop-Eleches and Urquiola 2011 for an example of how the impact of students winning lotteries to attend more selective schools is attenuated by school and household responses).

2. Model

The aim of this section is to offer an analytical framework to organize the empirical investigation and to understand the results. Becker and Tomes (1976) provide a classic model of the role of parents in spending on educational inputs, but do not model the interaction of school and household inputs. Todd and Wolpin (2003) allow for the possible substitutability of household and school inputs, but do not offer an explicit optimization model to derive empirical predictions. The contribution of our model is to specify the household's dynamic optimization

[3]Examples include school grants under the Sarva Shiksha Abhiyan (SSA) program in India, the Bantuan Operasional Sekolah (BOS) grants in Indonesia, and several similar school grant programs in African countries (see Reinikka and Svensson 2004 for descriptions of school grant programs in Uganda, Tanzania, and Ghana).

[4] Houtenville and Conway (2008) estimate an achievement production function that includes measures of parental effort and find that parental effort is negatively correlated with school resources.

[5] Of course, not all school inputs are substitutes. As we show in Section 2, these predictions do not hold for school inputs that are complementary to household inputs.

problem, solve it subject to both budget and production function constraints, and to derive the Euler equation that shows the optimal growth path of test scores (based on an appropriate shadow price of the cost of investing in educational inputs in each period).[6] We use this solution to discuss the differential impact of anticipated and unanticipated school inputs on test-score improvements and show how this varies based on whether school and household spending are substitutes or complements.

A household derives utility from the test scores of a child, *TS*, and the consumption of other goods, *X*. The household maximizes an inter-temporal utility function *U(.)*, additive over time and states of the world with discount rate β(<1), subject to an inter-temporal budget constraint. Finally, test scores are determined by a production function relating current achievement TS_t to past achievement TS_{t-1}, household educational inputs z_t, school inputs w_t, non time-varying child characteristics μ and non time-varying school characteristics η. We assume that household utility is additively separable, increasing and concave in test scores and other goods [A1]; and that the production function for test scores is given by $TS_t = F(TS_{t-1}, w_t, z_t, \mu, \eta)$ where $F(.)$ is concave in its arguments [A2].

Under [A1] and [A2] the household problem is

$$Max_{(X_t, z_t)}\, U_\tau = E_\tau \sum_{t=\tau}^{T} \beta^{t-\tau}[u(TS_t) + v(X_t)] \quad (1)$$

$$s.t. A_{t+1} = (1+r).(A_t + y_t - P_t X_t - z_t) \quad (2)$$

$$TS_t = F(TS_{t-1}, w_t, z_t, \mu, \eta) \quad (3)$$

$$A_{T+1} = 0 \quad (4)$$

Here u and v are concave in each of their arguments. The inter-temporal budget constraint, Equation (2), links asset levels A_{t+1} with initial assets A_t, private spending on educational inputs z_t, income y_t and the consumption of other goods, X_t. The price of educational inputs is the numéraire, the price of other consumption goods is P_t and r is the interest rate. The production function constraint, Equation (3) dictates how inputs are converted to educational outcomes, and the boundary condition, Equation (4) requires that at $t=T$, the household disposes of all remaining assets so that all loans are paid back and there is no bequest motive. We treat test scores as the observable measure of human capital. The latter is what parents care about, while

[6]This relates closely to the discussion on durable goods and inter-temporal household optimization; see Deaton and Muellbauer (1980), Jacoby and Skoufias (1997), Foster (1995) and Dercon and Krishnan (2000).

the former is what they observe and optimize with respect to.[7]

In this formulation, credit markets are perfect so that there are no bounds on A_{t+1} apart from Equation (4).[8] Moreover, households choose only the levels of X_t and z_t so that school inputs, w_t are beyond its control. In the contexts studied here, this is a reasonable assumption since school resources are allocated at state or federal levels and are not tied to a local property tax that residents may choose (unlike in the US). At the time the household makes its decision, it knows the underlying stochastic process governing w_t but not the actual level; we assume that school inputs are a source of uncertainty in the model—for simplicity, the only source.

Maximization of Equation (1) subject to Equations (2) and (3) provides a decision rule related to TS_t, characterizing the demand for test scores. Since test scores are a stock, we define a per-period price for test scores as the user-cost of increasing the stock in one period by one unit, i.e., the relevant (shadow) price in each period for the household. As in the durable goods literature (Deaton and Muellbauer 1980), the user cost, evaluated at period *t* is (see Das et al. (2004) for its derivation):

$$\pi_t = \frac{1}{F_{z_t}(.)} - \frac{F_{TS_t}(.)}{(1+r)F_{z_{t+1}}(.)} \qquad (5)$$

Here, the first term measures the cost of taking resources at *t* and transforming them into one extra unit of test scores. When implemented through a production function, the cost of buying an extra unit is the inverse of the marginal product of spending, $F_{z_t}(.)$. However, since *TS* is durable, increasing *TS* in period *t*, reduces the cost of acquiring *TS* in period *t+1* proportional to $F_{TS_t}(.)$ and the second term thus measures the present value of this reduction in cost in the next period, expressed in monetary terms.[9] Given the user cost, the first-order Euler condition

[7] The formulation can be seen as a short-cut for an alternative set up in which parents derive future utility from the flow of returns to the child's stock of human capital, as in a more standard human capital investment model. As we are mainly interested in deriving the optimal dynamic path for reaching the desired stock of human capital considering the costs and benefits of boosting test scores in current and future periods, the insights gained from using a human capital investment model are going to be similar, given the other assumptions made, especially the concavity of the period-by-period production function. Further, we assume that households care about the level of educational achievement, a stock. Results are unaffected if households care about the (instantaneous) flow from educational outcomes, provided that the flow is linear in the stock.

[8] It is straightforward to incorporate imperfect credit markets in this framework (see Das et al. 2004).

[9] In the durable goods literature, the user cost per period is derived by assuming that the good is sold in the second period. Though there is no "second hand market" for test scores, the shadow price for consuming a unit of test

determines the optimal path of educational outcomes between period *t-1* and *t* as:

$$E_{t-1}\left(\beta\frac{\pi_{t-1}}{\pi_t}\frac{\frac{\partial U}{\partial TS_t}}{\frac{\partial U}{\partial TS_{t-1}}}\right)=1 \qquad (6)$$

This is a standard Euler equation stating that along the optimal path, test scores will be smooth, so that the marginal utilities of educational outcomes will be equal in expectations, appropriately discounted and priced. Finally, the concavity of the production function *in each time period* will limit the willingness of households to boost education fast since the cost is increasing in household inputs.[10] Starting from low levels in childhood, the optimal path will be characterized by a gradual increase in educational achievement over time.

Under the further assumptions that household utility is additively separable and of the CRRA form, and that marginal utility is defined as $TS_t^{-\rho}$ (ρ the coefficient of relative risk aversion), Equation (6) can be rewritten as:

$$\left(\frac{TS_t}{TS_{t-1}}\right)^{-\rho}\frac{\beta\pi_{t-1}}{\pi_t}=1+e_t \qquad (7)$$

Where e_t is an expectation error, uncorrelated with information at *t-1*. Taking logs and expressed for child *i*, we obtain the optimal growth path:

$$ln\left(\frac{TS_{it}}{TS_{it-1}}\right)=\frac{1}{\rho}ln\,\beta-\frac{1}{\rho}ln(\frac{\pi_{it}}{\pi_{it-1}})+\frac{1}{\rho}ln(1+e_{it}) \qquad (8)$$

which is determined by the path of user-costs, and a term capturing surprises.

In this paper, we do not aim to use the structural dynamic model to estimate an impulse-response function over time to an unexpected change in inputs (the data requirements for that exercise are beyond any education data set we know of). However, we can use this theoretical model to derive an empirical model that nests some key predictions on how anticipated and unanticipated inputs affect the path of test scores. To derive these, assume that school resources are not known with certainty until households make decisions regarding their own inputs. Let

scores derived above is similar to those derived in the durable goods literature (see Foster 1995 for a similar derivation of the rental-equivalent price of boosting nutritional status in one period).

[10] The "per-period" concavity of the education production function can be motivated in several ways, the most intuitive of which is the existence of limits to how much a student can learn in a given period of time. While the unit of time is not specified in the model (as in the consumption smoothing literature in general), it is natural to consider the unit to be one year in the context of education, since decisions regarding education are typically made prior to the start of the school year.

$w_t^a\,(w_t^u)$ be inputs at time t that were anticipated (unanticipated) at *t-1*. For *unanticipated* increases in school inputs, households are unable to respond till the next time period and are therefore pushed off the optimal path (see footnote 10). The increase in educational achievement in period *t* is given by $F_{w_t} dw_t^u$, and the change in the growth path is given by $ln(TS_t + w_t^u F_w)$ which is strictly positive.

In the case of anticipated increases, the effect on the path of outcomes will depend on the impact on the user-cost of educational achievement at *t*, since there is no *direct* impact on the budget constraint at *t* (all information related to the anticipated inputs, including the budget constraint, will have been incorporated into the programming problem at *t-1*). Using the implicit function theorem with Equation (5) and assuming $TS_t = (1-\delta)TS_{t-1} + F(w_t, z_t, \mu, \eta)$ where the Hessian of $F(.)$ is negative semi-definite, we get

$$\frac{d\pi_t}{dw_t^a} = -\frac{F_{z_t w_t}}{F_{z_t}^2} \gtrless 0 \;\; if \;\; F_{z_t w_t} \lessgtr 0 \qquad (9)$$

The change in the optimal growth path is given by

$$\frac{\partial(\Delta_{t-1}^t lnTS)}{\partial w_t^a} = -\frac{1}{\rho}\left(\frac{\partial ln\pi_t}{\partial w_t}\right) \qquad (10)$$
$$= \frac{1}{\rho}\frac{1}{\pi_t}\frac{F_{zw}}{F_z^2} \gtrless 0 \;\; if \;\; F_{zw} \gtrless 0$$

If household and school inputs are technical substitutes so that $F_{z_t w_t} < 0$, anticipated increases in school inputs at *t* increase the relative user-cost of boosting *TS* at *t*, resulting in lower growth of test scores, ceteris paribus, between *t* and *t-1*.[11] Households have (price) incentives to shift resources for educational spending to *t-1*, boosting educational achievement at *t-1* in anticipation of the higher resources at *t*, and also to take advantage of the higher overall resources for educational inputs that allow them to spend relatively less on educational inputs

---

[11] In other words, if $F_{z_t w_t} < 0$, an increase in w_t will decrease the marginal product of z_t (and therefore increase the price of boosting *TS* by increasing z_t). We do not model the schools' choice of inputs to spend on, but if their objective function is to maximize *TS*, they should optimally allocate cash grants across different inputs and therefore account for the degree of substitution with households. One way to interpret these results is that schools are constrained in what they can do and are hence unable to spend this funding on inputs that could not be easily substituted for by parental resources. These constraints could arise either due to thin markets, explicit restrictions on the use of the grants (to hire teachers for instance), an inability to exploit scale economies (for instance, to improve infrastructure), or parental preferences expressed through school committees to spend on substitutable items).

compared to other commodities. Thus, the effect of an unanticipated change is higher than that of an anticipated change: household spending on educational inputs at *t* is unchanged, as households cannot move some of their spending to *t-1*, or to other commodities, as they could with anticipated increases of government spending.[12,13,14]

Assuming identical risk preferences, an empirical specification consistent with (8) is:

$$ln\left(\frac{TS_{it}}{TS_{it-1}}\right) = \alpha_o + \alpha_1 \, ln\, w_{it}^a + \alpha_2 \, ln\, w_{it}^u + \alpha_3 \Delta X_t + \varepsilon_{it} \qquad (11)$$

Here, w_{it}^a and w_{it}^u are anticipated and unanticipated changes in school inputs, measured in this paper by the flows of funds, while ΔX_t reflects all other sources of changes in the user cost between *t* and *t-1*. The core prediction is that the marginal effect of anticipated funds is lower than that of unanticipated funds when household and school inputs are substitutes. Finally, it is easy to see that if a portion of what the econometrician regards as unanticipated was anticipated by the household (or was substitutable even after the 'surprise' arrival of the school grant), then the estimate of α_2 will be a lower bound of the true production function effect.

3 The AP School Block Grant Experiment

3.1 Background and Context

Andhra Pradesh (AP) is the 5th largest state in India, with a population of over 80 million, with over 70% living in rural areas. AP is close to the all-India average on various measures of human development such as gross enrollment in primary school, literacy, and infant mortality, as

[12] If school and households inputs are technical complements, increasing school inputs at *t* will increase the marginal productivity of household inputs at t, and through the decline in user-costs lead to higher growth in test scores along the optimal path between t-1 and t. *Anticipated* lower user costs for educational inputs at t relative to t-1 create incentives to shift resources from t-1 to t, leading to a higher growth of test scores between t and t-1. Whether this reduces spending and therefore test scores at t-1 depends on preferences, as households have incentives to keep the optimal path of test scores smooth, while taking advantage of the additional government spending at t to spend more on other commodities.

[13] The model above is written as if there is only one type of school and household inputs. It is straightforward to allow for multiple inputs, taking *w* and *z* as *vectors* of educational inputs in the model. Different inputs could have different cross-derivatives, implying different degrees of technical substitutability, so that the extent to which the household may substitute for school spending on particular inputs may differ. We return to this issue in the conclusion, while discussing the broader implications of our results.

[14] Introducing household credit constraints does not fundamentally change these predictions, as anticipated increases in school spending will then alleviate the overall budget constraint of the household, and re-optimizing will imply that the gain in overall budget will be re-allocated across all commodities in the households' utility function (see Das et al. 2004).

well as on measures of service delivery such as teacher absence (Kremer et al. 2005). There are a total of over 60,000 government primary schools in AP and over 70% of children in rural AP attend government-run schools (Pratham 2010).

The average rural primary school is quite small, with total enrollment of around 80 to 100 students and an average of 3 teachers across grades one through five.[15] Teachers are well paid, with the average salary of regular civil-service teachers being over Rs. 8,000/month and total compensation including benefits being over Rs. 10,000/month (per capita income in AP is around Rs. 2,000/month). Regular teachers' salaries and benefits comprise over 90% of non-capital expenditure on primary education in AP, leaving relatively little funds for recurring non-teacher expenses.[16]

Some of these funds are used to provide schools with an annual grant of Rs. 2,000 for discretionary expenditures on school improvement and to provide each teacher with an annual grant of Rs. 500 for the purchase of classroom materials of the teachers' choice. The government also provides children with free text books through the school. However, compared to the annual spending on teacher salaries of over Rs. 300,000 per primary school (three teachers per school on average) the amount spent on learning materials is very small. It has been suggested therefore that the marginal returns to spending on learning materials used directly by children may be higher than more spending on teachers (Pritchett and Filmer 1999). The AP School Block Grant experiment was designed to evaluate the impact of providing schools with grants for learning materials, and the continuation of the experiment over two years (with the provision of a grant each year) allows us to test the differences between unanticipated and anticipated sources of school funds.

3.2 Sampling, Randomization, and Program Description

The school block grant (BG) program was evaluated as part of a larger education research initiative (across 500 schools) known as the Andhra Pradesh Randomized Evaluation Studies (AP RESt), with 100 schools being randomly assigned to each of four treatment and one control groups.[17] We sampled 5 districts across each of the 3 socio-cultural regions of AP in proportion

[15] This is a consequence of the priority placed on providing all children with access to a primary school within a distance of 1 kilometer from their homes.

[16] Funds for capital expenditure (school construction and maintenance) come from a different part of the budget. Note that all figures correspond to the years 2005 - 07, which is the time of the study, unless stated otherwise.

[17] The AP RESt is a partnership between the government of AP, the Azim Premji Foundation (a leading non-profit organization working to improve primary education in India), and the World Bank. The Azim Premji Foundation

to population. In each of the 5 districts, we randomly selected one administrative division and then randomly sampled 10 mandals (the lowest administrative tier) in the selected division. In each of the 50 mandals, we randomly sampled 10 schools using probability proportional to enrollment. Thus, the universe of 500 schools in the study was representative of the schooling conditions of the typical child attending a government-run primary school in rural AP.

The school year in AP starts in mid June, and baseline tests were conducted in the 500 sampled schools during late June and early July, 2005.[18] After the baseline tests were scored, 2 out of the 10 project schools in each mandal were randomly allocated to one of 5 cells (four treatments and one control). Since 50 mandals were chosen across 5 districts, there were a total of 100 schools (spread out across the state) in each cell. The geographic stratification allows us to estimate the impact of the program with mandal-level fixed effects and thereby net out any common factors at the lowest administrative level of government, and also improve the efficiency of the estimates of program impact.

Since no school received more than one treatment, we can analyze the impact of each program independently with respect to the control schools without worrying about any confounding interactions. The analysis in this paper is based on the 200 schools that comprise the 100 schools randomly chosen for the school block grant program and the 100 that were randomly assigned to the comparison group. Table 1 shows summary statistics of baseline school and student characteristics for both treatment and comparison schools and the null of equality across treatment groups cannot be rejected for any of the variables.[19]

As mentioned earlier, the block grant intervention targeted non-teacher and non-infrastructure inputs directly used by students. The block grant amount was set at Rs. 125 per student per year (around $3) so that the average additional spending per school was the same across all four programs evaluated under the AP RESt.[20] After the randomization was conducted, project staff from the Azim Premji Foundation (APF) personally went to selected

(APF) was the main implementing agency for the study. The details of the other interventions are provided in Muralidharan and Sundararaman (2010, 2011).

[18] The selected schools were informed by the government that an external assessment of learning would take place in this period, but there was no communication to any school about any of the treatments at this time.

[19] Table 1 shows sample balance between the comparison schools and those that received the block grant, which is the focus of the analysis in this paper. The randomization was done jointly across all treatments and the sample was also balanced on observables across the other treatments.

[20] The block grant was set on the basis of the number of students who took the baseline tests as opposed to the number of students enrolled (except for the first grade where there was no baseline test). This ensured that schools that inflated enrollment (which is not uncommon in India) were not rewarded with a larger grant.

schools to communicate the details of the school block grant program (in August 2005). The schools had the freedom to decide how to spend the block grant, subject to guidelines that required the money to be spent on inputs directly used by children. Schools receiving the block grant were given a few weeks to make a list of items they would like to procure. The list was approved by the project manager from APF, and the materials were jointly procured by the teachers and the APF field coordinators and provided to the schools by September, 2005. This method of grant disbursal allowed schools to choose inputs that they needed, but ensured that corruption was limited and that the materials reached the schools and children.

APF field coordinators also informed the schools that the program was likely to continue for a second year subject to government approval. Thus, while program continuation was not guaranteed, the expectation was that it was likely to continue for a second year. Schools were told early in the second year (June 2006) that they would continue being eligible for the school grant program and the same procedure was followed for disbursal of materials (no money was handed over to schools or teachers, and procurement was conducted jointly).

Table 2 shows that the majority of the grant money was spent on student stationery such as notebooks, and writing materials (over 40%), classroom materials such as charts (around 25%), and practice materials such as workbooks and exercise books (around 20%). A small amount (under 10%) of the grant was spent in the first year on student durable items like school bags, and plates/cups/spoons for the school mid-day meal program. This amount seems to have been transferred to stationery and writing materials in the second year. We also see that the overall spending pattern at the school level is quite stable across the first and second year of the grant. Many of these items could be provided directly by parents for their children, suggesting a high potential for substitution.

3.3 Data

Data on household expenditure on education was collected from a survey that attempted to cover every household with a child in a treatment or comparison school and administered a short questionnaire on education expenditures on the concerned child during the previous school year.[21] Data on household spending was collected at three points in time – alongside the baseline tests for spending incurred in the pre-baseline year (Y0), during the second year of the

[21] The data was collected from a short survey that was only based on the “main” child who was being covered in the school assessments and not for other siblings or other components of household spending.

program about spending during the first year (Y1), and after two full years of the program about spending during the second year (Y2). Data on household education spending was collected retrospectively to ensure that this reflected all spending during the school year.[22]

The data on learning outcomes used in this paper comprise of independent assessments in math and language (Telugu) conducted at the beginning of the study (June-July, 2005), and at the end of each of the two years of the experiment. For the rest of this paper, Year 0 (Y0) refers to the baseline tests in June-July 2005; Year 1 (Y1) refers to the tests conducted at the end of the first year of the program in March-April, 2006; and Year 2 (Y2) refers to the tests conducted at the end of the second year of the program in March-April, 2007. All analysis is carried out with normalized test scores, where individual test scores are converted to z-scores by normalizing them with respect to the distribution of scores in the control schools on the same test.

3.4 Results

3.4.1 Household Spending

We estimate:

$$\ln z_{ijkt} = \beta_0 \cdot Y_0 + \beta_1 \cdot Y_1 + \beta_2 \cdot Y_2 + \beta_3 \cdot BG \cdot Y_0 + \beta_4 \cdot BG \cdot Y_1 + \beta_5 \cdot BG \cdot Y_2 + \beta_m \cdot Z_m + \varepsilon_{ijk} \qquad (12)$$

where $ln\, z_{ijkt}$ is the expenditure incurred by the household on education of child *i*, at time *t* (*j*, *k*, denote the grade, and school), Y_n is the project year, and *BG* is an indicator for whether or not the child was in a "block grant" school. All regressions include a set of mandal-level dummies (Z_m) to account for stratification and to increase efficiency, and standard errors are clustered at the school level. The parameters of interest are β_3, which should equal zero if the randomization was valid (no differential spending by program households in the year prior to the intervention); β_4, which measures the extent to which household spending adjusted to an unanticipated increase in school resources (since the block grant program was a surprise in the first year of the project), and β_5, which measures the response of household spending to an anticipated increase in school resources (since the grant was mostly anticipated in the second year).[23]

[22] We obtained spending data from a total of 8,612 households for Y0 (no data was collected for retrospective spending on children in grade 1, because it was their first year in school), 13,572 households for Y1, and 10,189 households for Y2.

[23] We say "mostly anticipated" because it was not guaranteed that the program would be continued to the second year, but field reports suggest that the perception of the likelihood of continuation was high enough that households waited to see the materials provided by the schools before doing their own spending.

Table 3 confirms that β_3 and β_4 are not significantly different from zero while β_5 is significantly negative. We report the results both with and without a full set of household controls, and the results are unchanged. The estimated elasticity of -0.21 suggests that at the mean household expenditure for the comparison group (Rs 454 in Y2), the per-child grant of Rs. 125 would be substantially offset, and we cannot reject that the substitution is 100% (the point estimate of the offset is 76%).[24]

These findings are fully consistent with the predictions of the model: in Y1, households had limited ability to adjust to the unexpected grant; in Y2, household spending was able to adjust in anticipation of provision of materials by the school (using the grant). Evidence from field interviews suggests that the majority of household spending on education occurs at the start of the school year when notebooks, workbooks, stationery and writing materials are purchased. If an additional school grant arrives after this initial spending has taken place (as was the case in Y1) and is spent on additional learning materials by the school, households may not have been able to sell materials already purchased. However, once they knew about the school grant program, they would have been able to re-optimize their spending at the start of the next school year. Thus, the most likely mechanism for the results observed in Table 3 appears to be that the grant was unanticipated in the first year (and arrived after the majority of school spending for the year had taken place), but was anticipated in the second year in advance, which allowed households to re-optimize their own spending.

3.4.2 Student Test Scores

Our default specification for studying the impact of the school block grant, consistent with equation (11) uses the form:

$$\Delta T_{ijkm}(Y_n - Y_0) = \alpha + \gamma_j \cdot T_{ijkm}(Y_0) + \delta_n \cdot BG + \beta_m \cdot Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk} \qquad (13)$$

The main dependent variable of interest is $\Delta T_{ijkm}(Y_n - Y_0)$, which is the change in the normalized test score on the specific test (normalized with respect to the score distribution of the comparison schools) after *n* years of the program, where *i, j, k, m* denote the student, grade, school, and mandal respectively. Y_0 indicates the baseline tests, while Y_n indicates a test at the end of *n* years of the program. Including the normalized baseline test score on the right-hand side

[24] We use a logarithmic specification to be consistent with the model. We find identical results when estimating a linear model in levels of spending, and again cannot reject total substitution by households of the school grant in Y2.

improves efficiency due to the autocorrelation between test-scores across multiple periods.[25] These regressions include a set of mandal-level dummies (Z_m) and the standard errors are clustered at the school level. We also run the regressions with and without controls for household and school variables.

The *BG* variable is a school-level dummy indicating if the school was selected to receive the block grant (BG) program, and the parameter of interest is δ_n, which is the effect on normalized test score gains of being in a school that received the grant after *n* years. The random assignment of treatment ensures that the *BG* variable in the equation above is not correlated with the error term, and the estimate of the one-year and two-year treatment effects are therefore unbiased.[26]

At the end of the first year of the program, students in schools that received the block grant scored 0.09 and 0.08 standard deviations (SD) higher in mathematics and language (Telugu) than students in comparison schools, with both these differences being significant (Table 4 – columns 3 and 5). At the end of two years of the program, students in program schools scored 0.04 and 0.065 SD higher in mathematics and language, with neither of these effects being significant (Table 4 – columns 4 and 6). The addition of school and household controls does not significantly change the estimated value of δ_n, as would be expected given the random assignment of the grant program across schools (tables available on request).

We see that after two years of block grants, there is *no* significant effect on test scores, despite the gains after the first year. The size of gains after two years (with point estimates below the point estimates after Y1) suggest that the second year of block grants did not add much to learning outcomes, while decay of earlier gains may explain why average gains (in terms of point estimates) after Y2 are smaller than achieved after Y1. An alternate way of analyzing the data is to estimate a pooled regression of 1 and 2-year gains (T(Y1) –T(Y0), and T(Y2) – T(Y0)) as follows:

[25] The inclusion of the baseline test score also allows us to control also for individual heterogeneity correlated with baseline test-scores. The randomization ensures that the BG variable is uncorrelated with the error term. Since grade 1 children did not have a baseline test, we set the normalized baseline score to zero for these children (similarly for children in grade 2 at the end of two years of the treatment).

[26] We also check for differential post-treatment attrition of teachers and students and find that there is no differential attrition or turnover of teachers between "block grant" and "control" schools. However, there is a small amount of differential student participation in the test at the end of the first year of the program (with attrition from the baseline test-taking sample of 5.4% and 8.2% in the treatment and control groups respectively). As weaker students may drop out of the testing sample, this may bias our estimate of the first-year treatment effect downwards, but since the magnitude of differential attrition is small (2.8%), this bias is likely to be quite small, especially since baseline scores are controlled for. In the second year, however, there is no differential attendance on the end of year tests.

$$\Delta T_{ijkmt}(Y_n - Y_0) = \alpha_1 + \alpha_2 \cdot Y_2 + \gamma_j \cdot T_{ijkm}(Y_0) + \delta_1 \cdot BG + \delta_2 \cdot BG \cdot Y_2 + \beta_m \cdot Z_m + \varepsilon_{jk} + \varepsilon_k + \varepsilon_{ijk} \quad (14)$$

where δ_1 is the impact of the block grant program on test scores at the end of the first year, and δ_2 is the *additional* impact of the program in the second year. Table 5 shows these results and we see that δ_2 is always *negative* (though not significant) and we cannot reject that $\delta_1 + \delta_2 = 0$. The presence of decay (or fade out) of test scores[27] introduces a challenge for interpretation because δ_2 is the sum of the second year treatment effect and the decay of the first-year treatment effect (and these are not separately identified). However, the fact that the cumulative two-year effect is *lower* than the one-year effect (even though the grant was continued in the second year) strongly suggests that the school grant program did not lead to further improvement in Y2, and the negative estimates of δ_2 suggest decay in the gains from the first year.[28]

To shed further light on this issue, we present two estimates of the *second-year* treatment effect. First, we estimate equation (13) with the second-year gains (T(Y2) –T(Y1)) as the dependent variable controlling for Y1 scores. Note that this is not a consistent estimate of the second-year effect because Y1 scores are correlated with the treatment, and we cannot jointly estimate γ_j and δ_2. We therefore first estimate $\hat{\gamma}_j$ using only the control schools, and then estimate the following transformed version of (13):

$$\Delta T_{ijkm}(Y_2 - Y_1) - \hat{\gamma}_j \cdot T_{ijkm}(Y_1) = \alpha + \delta_2 \cdot BG + \beta_m \cdot Z_m + \varepsilon_k + \varepsilon_{jk} + \varepsilon_{ijk} \quad (15)$$

The results from estimating (15) suggest that the effect of the block grant program in the second year was close to zero in mathematics and 0.047 SD in Telugu, with a combined effect of only 0.02 SD - none of which are significant (Table 6 – Panel A).[29]

[27] Decay of test scores is well documented in the education literature. See Jacob and Lefgren (2009), Rothstein (2010), and Andrabi et al (2011) for more detailed discussions.

[28] Experimental evaluations of education interventions in developing countries find estimates of treatment effect decay in the range of 0.5 – 0.75 a year after the program is withdrawn (see Banerjee et al. 2007 and Glewwe et al. 2010 for examples). Assuming that the first-year treatment effects decay at a similar rate would yield an estimate of the second-year treatment effect between 0.01 and 0.03 SD (averaged across math and language).

[29] The key assumption required for (15) to be a consistent estimate of the second-year treatment effect is that $\hat{\gamma}_j$ be consistently estimated in a typical value-added model of test scores. While this is a standard assumption in the literature on test-score value addition, it need not hold true in general since measurement error of test scores would bias the estimate downwards, while unobserved heterogeneity in student learning rates would bias it upwards. However, Andrabi et al (2011) show that when both sources of biases are corrected for in data from Pakistan (which is a similar South Asian context), the corrected estimate is not significantly different from the OLS estimate used in the literature, suggesting that the bias in our approach is likely to be small.

Second, we estimate an average non-parametric treatment effect of the block grants in each year of the program by comparing the Y(n) scores for treatment and control students who *start at the same Y(n-1) score* (see the plots in Figures 1 and 2). The average non-parametric treatment effect (*ATE*) is the integral of the difference between the two plots, integrated over the density of the control school distribution, and is implemented as follows:

$$ATE = \frac{1}{100}\sum_{i=1}^{100}\left[\overline{T(Y_n(BG))} - \overline{T(Y_n(C))}\middle| T(Y_{n-1}(BG)), T(Y_{n-1}(C)) \in P_{i,n-1}(C)\right] \qquad (16)$$

where $P_{i,n-1}(C)$ is the *i*'th percentile of the distribution of *control school scores* in Y(n-1) and $T(Y_n(BG)), T(Y_n(C), T(Y_{n-1}(BG)), T(Y_{n-1}(C))$ are the test scores at the end of Y(n) and Y(n-1) in the treatment (BG) and control (C) schools respectively.

The intuition behind this estimate is straightforward. If test scores decay at a constant *rate*, then the absolute test score decay will be higher in the treatment schools in the second year (because test scores in these schools are higher after the first year), and the estimate of δ_2 in (14) will be an under-estimate of the second-year treatment effect. By matching treatment and control students on test scores at the end of Y1 and measuring the additional gains in Y2, we eliminate the role of decay because the treatment and control students being compared have the same Y1 score, and will therefore have the same absolute decay, and the difference in scores between these students at the end of Y2 will be an estimate of the second-year treatment effect that is not confounded by differential decay of test scores across treatment and control schools.[30] The treatment effects estimated at each percentile of the control school distribution are then integrated over the density of the *control distribution* to compute an average non-parametric treatment effect.[31]

[30] There are 2 main assumptions needed for this procedure to be valid. The first is that test scores decay at a constant rate and that the level of decay only depends on the current test score (and does not vary based on the inputs that produced these scores). This is a standard assumption in the estimation of education production functions (see Todd and Wolpin 2003). The second is that the effect of the treatment is the same at all points on the distribution of unobservables (since the treatment distribution is to the right of the control distribution after Y1, students who are matched on scores will typically not be matched on unobservables). While this assumption cannot be tested, it can be justified – especially after only one year of treatment, because a mean treatment effect of 0.085 SD after one year corresponds to only a 3.5 percentile shift at the median of a normal distribution. Also, as we note later, we find no heterogeneity of treatment effects along several observable school and household characteristics, suggesting that there may be limited heterogeneity in treatment effects across unobservables as well.

[31] Note that the treatment distribution in Y1 is to the right of the control distribution. Thus, integrating over the density of the control distribution adjusts for the fact that there are more students with higher Y1 scores in treatment schools and that test scores of these students will decay more (in absolute terms) than those with lower scores. In other words, treatment effects are calculated at every percentile of the control distribution and then averaged across

Figures 1 and 2 (which present average treatment effects across subjects) clearly suggest a positive treatment effect in Y1 and a much smaller effect in Y2. The average non-parametric treatment effect in the Y2 is close to zero, with point estimates of -0.03 SD for math and 0.05 SD for language, and the 95% confidence intervals on the treatment effects suggest that we cannot reject a zero treatment effect in the second year (Table 6 – Panel B). These estimates are almost identical to those in Panel A and again suggest that the second-year effect of the program on test scores (when the grants were anticipated) was close to zero.

Finally, we tested for heterogeneity of the block grant program effect across student and school characteristics by adding a set of characteristics and their interactions with the BG variable in (13).[32] The main result is the *lack* of heterogeneous treatment effects by several household and child-level characteristics, including household affluence. Even if we expect poor households to be more credit constrained and to be spending less than their desired 'optimal' amount of spending on education, this finding is not necessarily surprising: they have various needs, of which education is only one, and so we would not necessarily expect the poor to offset less of the grant than richer households.

The exception may be in cases where the value of the grant is higher than initial levels of household spending, since transaction costs in selling materials may make it difficult to fully monetize the value of the grant. The results suggest that even poor households were spending enough on education so as to substitute away most of the value of the school grant from their own spending. We verify this by looking at household spending on education in the control schools and find that only 12% of households report spending less than Rs. 125/year (the value of the grant) on their child's education, suggesting that the grant was infra-marginal for most households and could be offset easily.[33]

---

these percentiles *regardless of the number of treatment students* in each percentile of the control distribution at the end of Y1. Also, the estimate only uses students in the common support of the distribution of Y1 scores between treatment and control schools (less than 0.1% of students are dropped as a result of this).

[32] We tested the interaction of the program with school size, proximity to urban centers, school infrastructure, household affluence, parental literacy, caste, gender, and baseline test score.

[33] The non-parametric plots in Figures 1 and 2 provide a potentially richer picture of heterogeneity than that of a linear interaction, but the results are not significant. Figure 1 indicates a nearly uniform treatment effect in Y1 for students scoring in the bottom 70% of test scores, with a declining effect towards the top of the distribution. Figure 2 indicates that students scoring in the lowest 20% at the end of Y1 do slightly better in Y2 in the block grant schools. Since household affluence is strongly correlated with test scores, this figure suggests that even the anticipated grant may have increased total spending on materials for the poorest households (though we find no significant effect in the linear interaction discussed in the text).

3.5 Robustness

The evidence on household spending and test scores is consistent with a model in which households respond to anticipated school funding. The crowding out of private spending is sufficiently substantial to lead to no impact on test scores from anticipated school grants, while unanticipated changes positively impact the growth of test scores of children. We now consider the robustness of these results to alternative interpretations.

3.5.1 How different are parental expectations about the grant across the two years?

One possible concern with our approach may be that the distinction between anticipated and unanticipated funding is artificial, and households can similarly anticipate both sources after all. This is hard to sustain: as mentioned earlier, the schools had no reason whatsoever to expect the program in the first year, while the grant was eagerly anticipated by schools in the second year. Also, as suggested earlier, most household spending on education occurs at the start of the school year, whereas the announcement of the grant program was made around one and a half months into the school year in Y1 and materials were typically procured a few weeks after that. Thus, it is highly likely that materials bought with the grant supplemented the initial household spending and that the first-year program effect represents a "production function" effect of additional spending on school materials. In the second year of the program, field reports suggest that in many cases, parents were aware of the grant program, and waited to see what materials the school would buy with the grant before incurring their own expenditures on materials.[34] While, we do not explicitly measure or manipulate expectations, the discussion above suggests a clear difference in the degree of anticipation of funds in the first and second year.

3.5.2 What are the components of spending?

A further possible concern regarding our interpretation of the results is that it is possible that the grants in the first year were spent by schools on items that households cannot substitute for, while in the second year, the grants were spend on more substitutable items. We show that this is unlikely to be the case. Spending patterns across various categories are almost identical between the first and second years of the project and Table 2 clearly shows that the funds were spent on the same type of inputs both when they were unanticipated (first year) and anticipated (second year). This also helps rule out explanations based on diminishing returns to the items procured or

---

[34] This interpretation is further corroborated by field reports from household interviews after the program was withdrawn, which suggest that around two months into the school year, most parents had not bought the materials that they thought would be provided by the school

the durable nature of school materials. It is possible that some of the classroom materials purchased may be durable, and the results reflect diminishing returns to durables in the second year. However, we see that the same fraction of the grant was spent on classroom materials in both years, suggesting that even these materials needed to be replenished. We also explicitly record spending on durables (bags, uniforms, plates, etc.) and find that these accounted for less than 10% of spending in the first year, and under 1% in the second year.

3.5.3 Storage and Smoothing

In interpreting our results, a question that arises is whether households or schools could have smoothed the unexpected grant by either saving some of the funds or storing some materials for use in later years (if the materials had already been bought). On the school side, the program design did not provide schools the option of saving funds. They could have saved materials, but they spend on the same sets of materials in both years suggesting that storage was limited, and that the grant led to a near one for one increase in learning materials in the first year.

On the household side, we see that they do not reduce their expenditure in response to the unanticipated grant, but cannot fully rule out the possibility of some storage. But even if some smoothing via savings, storage, or durable goods spending by the school may have been possible, the coefficient on the unexpected grant is a *lower bound* on the production function parameter (because in this case, the full value of the grant will not have been spent in the same time period) and our results show that the production function effect of the school grant is positive – which would not have been apparent if the relationship between school grants and test scores were to have been estimated using anticipated grants (as we will see again in the Zambia results in section 4).

3.5.4 Other Budgetary Offsets

A further concern is the possibility that anticipated funds are offset by a reduction of other transfers to the program schools, which may explain the drop off in the second-year test scores in the treatment group. We rule this out by measuring the total grants received by the schools from all other sources and finding that there is no difference in year to year receipts of funds in either treatment or control schools. There is also no significant difference between the amounts received in treatment and control schools in any year, or a significant difference between any of these differences across the years (tables available on request).

3.5.5 Are parents behaving rationally?

Our results may raise the concern that parents are 'leaving human capital on the table' and not behaving rationally (as implied by the model). Specifically, if test scores can be increased by 0.09 SD by simply spending an extra $3/year (as indicated by the Y1 results), is it rational for parents to cut back their own spending in response to the grant and forego these gains to test scores? The data suggest that parents are not behaving irrationally, and that the extent of the offset yields an estimate of income elasticity of education spending between 1.8 and 4.6, which suggests that parents spend a greater share of income on education as income increases.[35] However, since the grant is fungible when provided in the form of books and materials, it is rational for households to offset a considerable fraction of the value of the grant (but not all of it) and to accept a correspondingly lower impact on test scores than when all the additional income was spent on education (as was the case in Y1). This impact may be positive, but is not significantly different from zero in our data.[36]

3.5.6 Time-varying teacher and parental effort

A final possibility we consider is that there is no direct link between the increased resources, the corresponding household responses, and the test-scores findings, but that the test-score findings are in fact mediated by some other process. One possible narrative could be that the test score response in Y1 is not a production function effect of the grant, but is instead due to increased teacher and school effort in response to receiving a 'gift' (as in the gift-exchange

[35] The income elasticity of education spending is calculated using 2007 data on total household spending (Rs. 8,738/year), and per capita government spending on education in Andhra Pradesh (Rs. 867/year) provided in Tilak (2009), and data on household spending on education from our surveys (Rs. 454/year in the control schools). Our point estimates suggest that 76% of the grant was offset in the second year (or that 24% of the increase in income was spent on education). If we include the government spending on education as part of the base spending on education, we get an education budget share of 13.7% and an estimated income elasticity of education spending of 1.8. If we do not include the government spending and only consider out of pocket expenditure, we get an education budget share of 5.1% and an estimated income elasticity of education spending of 4.6. These numbers are only suggestive, because there is an additional caveat in using the Y2 spending data to calculate the income elasticity of education spending. Specifically, note that being forced to spend all or most of the Y1 grant on education may have caused the household to overspend on education and as a result they may reduce spending in Y2 by more than they would in the case a general income increase (to return to their optimal long-term growth path of test scores). We also cannot measure the value of any materials that the household may have stored between Y1 and Y2. However, both these points would imply a higher income elasticity of education spending than our estimates above and do not detract from our main point, which is that households do not appear to be irrationally reducing their own spending.

[36] While we could not detect a significant impact in test-scores Y2, despite some increase in *overall* spending on education by schools and households, our best estimate of the gain in Y2, the point estimate in column [1] in table 6, suggests a gain of 0.0.02 SD, or a quarter of the gain in Y1, consistent with a net spending gain in Y2 that was a quarter of that in Y1. As we do not have data on what the households did with the extra cash, we cannot say much about the welfare impact of the program. However, since the grant was small, it would be difficult to detect significant increases in any particular component of household spending even with more detailed household surveys.

model of Akerlof 1982), whereas in Y2, the schools and teachers get 'habituated' to the grant, and then parents reduce spending while teachers reduce effort (see Gneezy and List 2006 for an example of this). Thus, the results could be driven by school-level factors as opposed to the differential household responses.

The data suggests that this is unlikely to be the case. We find no difference in teacher absence or teaching activity across treatment and control groups in either Y1 or Y2 or within treatment schools across Y1 and Y2. Furthermore, if such a 'gift exchange' idea was empirically relevant, we should expect similar patterns to be present in the other experiments conducted in the same setting, with considerably higher impact when programs start, but then dropping off to no impact when schools get habituated to the programs. We find that this is not the case. In schools provided with an extra contract teacher or with performance-linked pay for teachers (see Muralidharan and Sundararaman 2010 and 2011 for details), the 2-year effect is larger than the 1-year effect (and we cannot reject that the 2-year effect is twice the 1-year effect), and the block grant program is the only one where the 2-year effect is lower than the 1-year effect. Since teacher inputs (headcount or effort), cannot easily be substituted for by illiterate parents (while materials can), these results offer further support to our contention that the test score results in this paper most likely reflect the difference between a situation where households have not yet re-optimized their spending (Y1) and one where they have (Y2).

Overall, the considerable crowding out as found in the household spending analysis (Table 3) continues to offer a consistent, plausible, and parsimonious mechanism to explain our findings that test scores are significantly higher in program schools at the end of Y1, but not different across treatment and controls schools at the end of Y2.

A key question in considering the broader relevance of our results is the extent to which they can be replicated in other settings. Our data from Zambia allows us to test the main predictions of the model in a completely different context, and provide two additional advantages beyond external validity. First, the data come from a nationally scaled-up school grant program implemented by the Government of Zambia as a 'steady state' policy, and these results may be more directly relevant to other policy settings. The second advantage is that in addition to the predictable school grant, we also have data on a much more idiosyncratic source of school funding, which allows us to test the impact of both unanticipated and anticipated grants on test scores contemporaneously (whereas it was sequential in AP).

4 Zambia

4.1 Background and Context

The education system in Zambia is based on public schools (less than 2 percent of all schools are privately run) and the country has a history of high primary enrollment rates. Teacher salaries are paid directly by the central government, and account for the majority of spending on school-level resources; schools receive few other resources from the government. Parental involvement in schools is high with parents traditionally expected to contribute considerably to the finances of the school via fees paid through the Parent Teacher Association (PTA). Limited direct government funding for non-salary purposes during economic decline put pressure on parents to provide for inputs more usually provided by government expenditure. This customary arrangement regarding PTA fees changed in 2001; following an agenda of free education, all institutionalized parental contributions to schools, including formal PTA fees were banned in April 2001.

At the same time (in 2001), a rule-based cash grant through the government's Basic Education Sub-Sector Investment Program (BESSIP) was provided to every school to reverse some of the pressure on school finances arising from the banning of PTA fees. These grants were fixed at $600 per school ($650 in the case of schools with Grades 8 and 9) irrespective of school enrollment to exclude any discretion by the administration. The grant was managed via a separate funding stream from any other financial flows, and directly delivered to the school, via the headmaster. Spending decisions were made at the Annual General Meeting, before the start of the school year. The share of the BESSIP grant in overall school funding was considerable: for 76% of schools it was the only public funding for non-salary inputs, while its average share in total school resources was 86%.

The scheme also attracted much publicity, which increased its transparency. Combined with the simplicity of the allocation rule, this ensured that the grants reached their intended recipients. Disbursement was fast and reliable and 95 percent of all schools had received the stipulated amounts by the time of the survey and the remainder within 1 month of survey completion (Das et al. 2003).[37] Therefore, we expect that in the year of the survey (2002) the fixed cash grants would be anticipated by households making their education investment decisions for the year.

[37] This contrasts with the early experience in Uganda (Reinnika and Svensson 2004).

Furthermore, because the grants were fixed in size, there was considerable variation across schools in per-student terms due to underlying differences in enrollment.[38]

In addition to these predictable rule-based grants, districts also received some discretionary funding for non-salary purposes from the central government and aid programs. However, since the 1990s, these sources were highly unreliable and unpredictable, partly due to the operation of a "cash budget" in view of the poor macroeconomic situation, and partly due to the irregularity of much of the aid flows to the education sector (Dinh, et al. 2002). In 2002, the year of our survey, less than 24 percent of all schools received such discretionary grants and conditional on receipt, there was considerable variation with some schools receiving 30 times as much as others.[39] Conversations with district-level officials suggested that it was very difficult for schools to predict whether these grants would be received (and if so how much),[40] and as we discuss further below, there does not appear to be any correlation between receipt of these discretionary funds and observable characteristics of the schools. Overall, the share of discretionary resources was only about a tenth of the share of the teacher salary bill. Finally, few resources were distributed in kind to schools during the year of the survey (see Das et. al 2003).

This variation in the per-student rule-based grants as well as the variation in the receipt of discretionary funds allows us to study the impact of anticipated and unanticipated school grants on test score gains as discussed below

4.2 Sampling and Data

We collected data in 2002 from 172 schools in 4 provinces of Zambia (covering 58 percent of the population), where the schools were sampled to ensure that every enrolled child had an equal probability of inclusion. The school surveys provide basic information on school materials and funding as well as test scores for mathematics and English for a sample of 20 students in grade 5 in every school, who were tested in 2001 as part of an independent study and were then retested in 2002 to form a panel.

---

[38] The mean transfer per pupil was about $1.2, and the range of the 10th to the 90th percentile of the per-pupil grant was $0.3 to $2.5 confirming the wide variation in the grant amount.

[39] The average discretionary transfer per pupil in the sample was about $2.4; conditional on receiving it, this is about $9.8 per pupil.

[40] The most commonly told anecdote by district officials was that principals who happened to be visiting the district office at a time when some discretionary funds were available were the ones who were likely to obtain some of these funds for their schools. Similarly, conversations with principals indicated that they would often have a list of requests for funds when they would visit the district office but that it was quite unpredictable as to whether funds would be available.

A key advance over the existing literature on the impact of school spending on test scores is our ability to create a matched data set of spending between schools and households. We do this by collecting education expenditure data from 540 households matched to a sub-sample of 34 schools identified as "remote" using GIS mapping tools (defined as schools where the closest neighboring school was at least 5 kilometers away). From these schools, the closest village was chosen and 15 households were randomly chosen from households with at least one child of school-going age. The restriction of the household survey sample to 34 remote schools allows us to match household and school inputs in an environment where complications arising from school choice are eliminated. We use the entire sample of 172 schools to estimate the relationship between test scores and cash grants to schools (rule-based and discretionary). We use the sub-sample of 34 schools matched to 540 households to estimate the relationship between rule-based cash grants to schools and household expenditures on education.

4.3 Impact of School Grants on Test Scores

We explore the impact of different types of school grants using Equation (17), based on (11), modeling changes in standardized test-scores *TS* between *t* and *t-1* regressed on anticipated and unanticipated school funds, and a set of controls at *t-1* (capturing sources of heterogeneity):

$$\Delta TS_{it} = \alpha_o + \alpha_1 \ln w_{it}^a + \alpha_2 \ln w_{it}^u + \alpha_3 X_{t-1} + \varepsilon_{it} \qquad (17)$$

In (17), w_j^a and w_j^u are respectively anticipated (from the rule-based BESSIP grant) and unanticipated (from district-level discretionary sources) grants per student in school *j*, and X_{t-1} are a set of geographic and school level control variables.[41] The prediction is that $\alpha_1 < \alpha_2$: unanticipated spending will have a larger effect on test scores than anticipated spending.

We first present results from estimating equation (17) with only the anticipated grant, and our main result is that there is no correlation between variation in per-student rule-based school grants and test score gains. We then add an indicator for receipt of discretionary funds (that we argued earlier are difficult to anticipate relative to the rule-based grants) to estimate equation (17) and test $\alpha_1 < \alpha_2$. Recall that there is high variability in discretionary funding, with less than a quarter of the school sample receiving any funds, and high variance among schools receiving funds. We therefore present two functional forms - first with an indicator for receipt of any discretionary funds as a binary variable and second with a continuous measure for the amount of

[41] Geographic controls include province and rural/urban indicators; school controls include school-level variables such as related to characteristics of the head-teacher, the head of the Parent-Teacher Association and PTA fees.

discretionary funds received, including both linear and quadratic terms.

The first main result we see that there is no correlation between variation in rule-based, anticipated school grants and test score gains (Table 7 – columns 1 and 4). These results are similar to those observed in several other contexts and would suggest that "spending does not matter" for education outcomes. However, when we add an indicator for whether a school received discretionary funds (that we argued are difficult to anticipate), we find that students in schools receiving discretionary funds (with a median value of $3/student) gain an additional 0.10 SD in both English and Mathematics test scores (columns 2 and 5). When the discretionary funds are coded as a continuous variable, we find significant positive effects on English scores, but do not find any effect on Math scores (columns 3 and 6).[42]

One key threat to identification in the results above is the possibility that the discretionary/unanticipated grants may have been targeted to areas with the most potential improvement in test scores. Alternatively, parents and communities that cared enough to obtain these funds for their schools may also be motivated to increase test scores in other ways. We address this concern by comparing the characteristics of schools that do and do not receive these discretionary funds and find that there is no significant difference between these types of schools (Table 8 - Column 3). We also test if these observable characteristics can jointly predict whether a school would have received discretionary funds, and reject the joint significance of these characteristics (Table 8 - Column 4). While we cannot fully rule out omitted variable concerns, there is no evidence of differences between schools that do and do not receive these discretionary funds on observable characteristics.

To summarize, we find that variation in rule-based, well publicized source of funding are not correlated with test score gains, while less predictable funding sources are. These results highlight the potentially different impacts of unanticipated and anticipated school funds on test score gains, and the importance of making this distinction for empirical work. The second novel contribution of the empirical work in Zambia to the literature on the impact of school spending on test scores is our ability to analyze matched data on household and school spending, and study the possibility of household spending offsets as a possible mechanism for the lack of correlation between predictable grants and test score gains.

---

[42] Nonparametric investigation of the relationship between levels of discretionary funds and test score gains suggested a positive, but highly non-linear relationship for both English and Mathematics.

4.4.2 Household Spending

We estimate a cross-sectional household expenditure model for the 1195 children (from 540 households) matched to 34 schools in which household spending on school-related inputs is regressed on anticipated and unanticipated grants with and without a set of controls for child, household and school-level variables. We estimate:

$$\ln z_{ij} = \alpha + \beta_1 A_i + \beta_2 \ln w_j^a + \beta_3 \ln w_j^u + \beta_4 X_i + \varepsilon_i + \varepsilon_j \qquad (18)$$

where z_{ij} is the spending by the household on child i enrolled in school j, w_j^a and w_j^u are respectively anticipated (rule-based) and unanticipated (discretionary) grants per student in school j that matches to child i, and X_i are other characteristics of child i including assets owned by the household. We test $\beta_2 < \beta_3 = 0$, i.e., households respond negatively to the pre-announced, anticipated rule-based grants at the school level by cutting back their own funding, but are unable to respond to cash grants that are unanticipated.

We first present OLS results of estimating (18) without and with controls (Table 9 - Columns 1 and 2). However, one concern with OLS could be that w_j^a captures unobserved components of household demand operating through an enrollment channel (since the per-child rule-based grant will be smaller in schools with a larger enrollment). We therefore use the size of the eligible cohort in the catchment area as an instrument for school enrollment and therefore the level of per-student cash grants (columns 3 and 4).[43] This instrumentation strategy is similar to Case and Deaton (1999), and Urquiola (2006) in the case of class-size and more recently by Boonperm et al. (2009) and Kaboski and Townsend (forthcoming) in the context of large fixed grants to villages in Thailand. Using the size of the eligible cohort as an instrument for enrollment is especially credible in this context since we use only a sample of *remote* schools and can abstract away from issues of school choice. We also confirm that there is no correlation between the instrument and X_i.[44]

The results are consistent with the predictions from our model: across all specifications - OLS and IV - the estimated elasticity of substitution for anticipated grants ($\hat{\beta}_2$) is always

[43] The size of the catchment area is defined here as to the total number of children of the relevant age group in the five villages identified by the school as the most important sources of pupils.
[44] We can reject the hypothesis that the instrument is weak: the F-statistic of the first stage regression is above 10. The impact of an extra child in the catchment area on enrollment is 0.68 – which is close to the actual enrollment of about 80% in the sample.

negative and significant and ranges from -0.72 to -1.12 while the coefficient of unanticipated grants ($\hat{\beta}_3$) is small and insignificant. Crowding out appears large, and evaluated at the mean we cannot even reject the hypothesis that for each dollar spent on the rule-based grant per student, households reduce school expenditure by one dollar, while there is no substitution of discretionary, unanticipated spending. However, we place less emphasis on the latter result because only 4 out of the 34 remote schools (where we have household spending data) reported receiving any of the discretionary funds (whereas all 34 schools received the rule-based grant).

One concern with the result on spending offsets is that households in larger villages (which have smaller per capita anticipated funding) could have a different overall demand for education. We address this concern by comparing household expenditure across schools with different levels of rule-based grants. We divide schools into two categories - those receiving less than the median per-child rule-based grant ("low rule-based grant schools) and those receiving more than the median ("high rule-based grant schools) - and Table 10 shows school and household expenditure for these two types of schools. As expected from the definition, we find that the per-student grant is significantly lower in the "low rule-based grant" schools. However, household spending on education is significantly higher in these schools. Most importantly, there is *no significant difference* in total expenditure per child across these two school types. This suggests that overall demand for education is similar across the households in the sample, and that they compensate/offset for lower/higher spending at the school level.

4.5 Limitations and Robustness

The main caveat to the test score results in Zambia is the possibility that the discretionary funds are correlated with unobservables that could be correlated with test score gains, and the main caveat to the spending results is the possibility that households in larger villages have a different demand for education. While we cannot completely rule out these possibilities, Tables 8 and 10 suggest that these concerns may not be first order ones.

Other caveats seem less important. While we cannot attribute school spending to specific sources of funding (discretionary vs. rule-based), much spending at the school level from both sources appears to be substitutable. The total shares spent on those items most suitable for substitution (books, chalks, and stationery) add up to 57% and 47% respectively for schools without and with discretionary funding, suggesting that in both cases, substantial and similar spending occurs on items that could be substituted by households.

It is also hard to prove whether the discretionary spending was a true surprise, but the uncertainty related to the cash-budget meant that actual spending and budgets were far apart. The typical arrival of these funds at varying points during the school year suggest that households were unlikely to be able to respond to these (as suggested by the positive test score gains in these schools in Table 7, and the findings in Table 9). In addition, we see clearly in Table 10 that households do respond substantially to variations in the rule-based grants and that they spend much more/less in schools with lower/higher per-student rule-based funding. These different types of funding were also not used to offset each other, confounding the results. In fact, we find a *positive* but insignificant relationship between rule-based funding and discretionary funding [p-value=0.22].

Mirroring the results from the experimental design in AP, but this time from a nationwide program of school grants, the findings from Zambia suggest that the crowding out of household spending in response to a predictable stream of school funds is likely to be an important mechanism behind the lack of correlation between variation in anticipated school spending and test scores. While we cannot allay all possible identification concerns with cross-sectional evidence, the correlations presented are consistent with the model, and the model in turn provides a parsimonious and consistent framework to interpret the evidence.

5. Conclusion

Data on test-scores and household expenditures in the context of an experimental school grant program in Andhra Pradesh in India suggest that households reduce private educational spending in response to anticipated school grants. Consequently, school grants that are fully anticipated only have a limited impact on test-scores. Unanticipated grants elicit no household responses and do have positive impacts on learning. Cross-sectional data from a nationally-scaled up school grant program, matched with data on household spending and student test scores in Zambia show correlations that are consistent with the same interpretation. Finding the same result in different countries on different continents, with different implementing agencies, and in both experimental as well as scaled up programs suggests that the issue of household crowd out in the context of public education spending is likely to be of general relevance for both education research and policy.

These results have implications for common estimation techniques in the education production function literature, where achievement (or changes in achievement) is regressed on school inputs. Following Todd and Wolpin (2003), these estimates represent the policy effect of school inputs that combines both the effect of inputs on test scores through the production function, as well as household responses to such inputs. Our use of unanticipated inputs allows the examination of both effects separately, thus shedding more light on the process through which school inputs may or may not affect educational attainments.

This distinction between anticipated and unanticipated inputs could account for the wide variation in estimated coefficients of school inputs on test scores (Glewwe 2002, Hanushek 2003, or Kreuger 2003). The production function framework does not separate anticipated from unanticipated inputs and so the regressor is a combination of these two different variables. The estimated coefficient is bounded below by the policy effect and above by the production function parameter; the distance from either bound depends on the extent to which the schooling inputs were anticipated, and the extent to which they were substitutable by households. While experimental evaluations of education interventions typically overcome selection and omitted variable concerns, the distinction highlighted in this paper is relevant even for experiments, since the interpretation of experimental coefficients depends on the time horizon of the evaluation and whether this was long enough for other agents (especially households) to re-optimize their own inputs.

Although we find evidence of high crowding out of anticipated inputs, our results do not suggest an education policy where inputs are provided unexpectedly. Although test scores in the current period increase with unanticipated inputs, the additional consumption will push households off the optimal path. In subsequent periods, therefore, they will readjust expenditures until the first-order conditions are valid again – unanticipated inputs in the current period will not have persistent effects in the future (except due to the durable nature of some inputs). The policy framework that is suggested under this approach involves a deeper understanding of the relationship between public and private spending, acknowledging that this may vary across different components of public spending. Thus, a policy implication of our results is that schooling inputs that are less likely to be substituted away by households may be better candidates for government provision.[45]

What might such inputs be? One important example may be teaching inputs, whereby the combination of economies of scale in production (relative to private tuition), difficulty of substituting for teacher time by poorly educated parents, or the generic non-availability of trained personnel in every village could make public provision more efficient (see Andrabi et al., 2010). In a parallel experiment on the provision of an extra teacher to randomly-selected schools in Andhra Pradesh, Muralidharan and Sundararaman (2010) find that the impact of the extra teacher was identical in both the first and second year of the project – suggesting that teacher inputs were less likely to be substituted away. Similarly, inputs like school infrastructure that retain some aspects of public-goods and would thus be under-provided by non-coordinating households may be a good candidate for government provision.

The approach followed here of treating test scores as a household maximization problem, with the production function acting as a constraint, explicitly recognizes the centrality of households in the domain of child learning, with important implications for both estimation and policy. These issues go beyond the study of the impact of public expenditures on education, but apply similarly to other areas of public spending, such as health and anti-poverty programs. More broadly, analysis of the impact of development programs in general will benefit from paying careful attention to the behavioral responses of households to enrich our understanding of observed variation in policy impacts in different settings and over different time horizons.

[45] An alternative could be to give very large grants to school. The anticipated grant in both countries was relatively small. For example, in AP only 12% of households were spending less than the per pupil school grant. If a grant larger than household spending had been given, then crowding out of household spending would have been bounded (assuming that they could not sell the inputs provided by the school), and the additional school grant may have had a positive impact on test-scores as total spending by schools and households would have increased.

References

AKERLOF, G. (1982): "Labor Contracts as Partial Gift Exchange," *Quarterly Journal of Economics*, 97, 543-569.

ANDRABI, T., J. DAS, and A. KHWAJA (2010): "Students Today, Teachers Tomorrow? Identifying Constraints on the Provision of Education," Harvard University.

ANDRABI, T., J. DAS, A. KHWAJA, and T. ZAJONC (2011): "Do Value-Added Estimates Add Value? Accounting for Learning Dynamics," *American Economic Journal: Applied Economics*, 3, 29-54.

AUTOR, D. H., and M. G. DUGGAN (2003): "The Rise in the Disability Rolls and the Decline in Unemployment," *The Quarterly Journal of Economics*, 118, 157-205.

BANERJEE, A., S. COLE, E. DUFLO, and L. LINDEN (2007): "Remedying Education: Evidence from Two Randomized Experiments in India," *Quarterly Journal of Economics*, 122, 1235-1264.

BECKER, G. S., and N. TOMES (1976): "Child Endowments and the Quantity and Quality of Children," *Journal of Political Economy*, 84, S143-S162.

BOONPERM, J., J. HAUGHTON, and S. R. KHANDKER (2009): "Does the Village Fund Matter in Thailand?," The World Bank.

CASE, A., and A. DEATON (1999): "School Inputs and Educational Outcomes in South Africa," *Quarterly Journal of Economics*, 114, F1047-F84.

CUTLER, D. M., and J. GRUBER (1996): "The Effect of Medicaid Expansions on Public Insurance, Private Insurance, and Redistribution," *The American Economic Review*, 86, 378-383.

DAS, J., S. DERCON, J. HABYARIMANA, and P. KRISHNAN (2003): "Rules Vs. Discretion: Public and Private Funding in Zambian Education," Development Research Group, The World Bank.

— (2004): "When Can School Inputs Improve Test Scores?," World Bank Policy Research Working Paper 3217.

DEATON, A., and J. MUELLBAUER (1980): *Economics and Consumer Behavior*. Cambridge, UK: Cambridge University Press.

DERCON, S., and P. KRISHNAN (2000): "In Sickness and in Health: Risk Sharing within Rural Households in Rural Ethiopia," *Journal of Political Economy*, 108, 688-727.

DINH, H. T., A. ADUGNA, and B. MYERS (2002): "The Impact of Cash Budgets on Poverty Reduction in Zambia: A Case Study of the Conflict between Well-Intentioned Macroeconomic Policy and Service Delivery to the Poor," The World Bank.

EISSA, N., and J. B. LIEBMAN (1996): "Labor Supply Response to the Earned Income Tax Credit," *The Quarterly Journal of Economics*, 111, 605-637.

FOSTER, A. D. (1995): "Prices, Credit Markets and Child Growth in Low-Income Rural Areas," *The Economic Journal*, 105, 551-570.

GLEWWE, P. (2002): "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes " *Journal of Economic Literature*, 40, 436-482.

GLEWWE, P., N. ILIAS, and M. KREMER (2010): "Teacher Incentives," *American Economic Journal: Applied Economics*, 2, 205-227.

GLEWWE, P., and M. KREMER (2006): "Schools, Teachers, and Education Outcomes in Developing Countries," in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch: North-Holland.

GNEEZY, U., and J. LIST (2006): "Putting Behavioral Economics to Work: Field Evidence on Gift Exchange," *Econometrica*, 74, 1365-1384.

HANUSHEK, E., and J. A. LUQUE (2003): "Efficiency and Equity in Schools around the World," *Economics of Education Review*, 20, 481-502.

HANUSHEK, E. A. (2002): "Publicly Provided Education," in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. S. Feldstein. Amsterdam: North-Holland, 2045-2141.

— (2003): "The Failure of Input-Based Schooling Policies," *Economic Journal*, 113, F64-98.

HOUTENVILLE, A. J., and K. S. CONWAY (2008): "Parental Effort, School Resources, and Student Achievement," *Journal of Human Resources*, 43, 437-453.

JACOBY, H. G., and E. SKOUFIAS (1997): "Risk, Financial Markets, and Human Capital in a Developing Country," *The Review of Economic Studies*, 64, 311-335.

KABOSKI, J. P., and R. M. TOWNSEND (Forthcoming): "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*.

KANYIKA, J., C. T. SAKALA, M. MWALE, R. MUSUKU, G. G. MWEEMBA, and T. NAKASWE-MUSAKANYA (2005): "Learning Achievement at the Middle Basic Level: Zambia's National Assessment Survey Report."

KREMER, M., K. MURALIDHARAN, N. CHAUDHURY, F. H. ROGERS, and J. HAMMER (2005): "Teacher Absence in India: A Snapshot," *Journal of the European Economic Association*, 3, 658-67.

KRUEGER, A. (2003): "Economic Considerations and Class Size," *Economic Journal*, 113, 34-63.

MEYER, B. D. (1990): "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58, 757-782.

MOFFITT, R. A. (2002): "Welfare Programs and Labor Supply," in *Handbook of Public Economics*: Elsevier B.V., 2393-2430.

MURALIDHARAN, K., and V. SUNDARARAMAN (2010): "Contract Teachers: Experimental Evidence from India," UC San Diego.

— (2011): "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 119, 39-77.

POP-ELECHES, C., and M. URQUIOLA (2011): "Going to a Better School: Effects and Behavioral Responses," Columbia University.

PRATHAM (2010): *Annual Status of Education Report*.

PRITCHETT, L., and D. FILMER (1999): "What Education Production Functions Really Show: A Positive Theory of Education Expenditures," *Economics of Education Review*, 18, 223-39.

REINIKKA, R., and J. SVENSSON (2004): "Local Capture: Evidence from a Central Government Transfer Program in Uganda," *The Quarterly Journal of Economics*, 119, 679-705.

TILAK, J. B. G. (2009): "Household Expenditure on Education and Implications for Redefining the Poverty Line in India," Planning Commission of India.

TODD, P. E., and K. I. WOLPIN (2003): "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal*, 113, F3-33.

URQUIOLA, M. (2006): "Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia," *Review of Economics and Statistics*, 88, 171-177.

URQUIOLA, M., and E. VERHOOGEN (2009): "Class-Size Caps, Sorting, and the Regression-Discontinuity Design " *The American Economic Review*, 99, 179-215.

Figure 1: Average non-parametric treatment effects (Y1 on Y0)

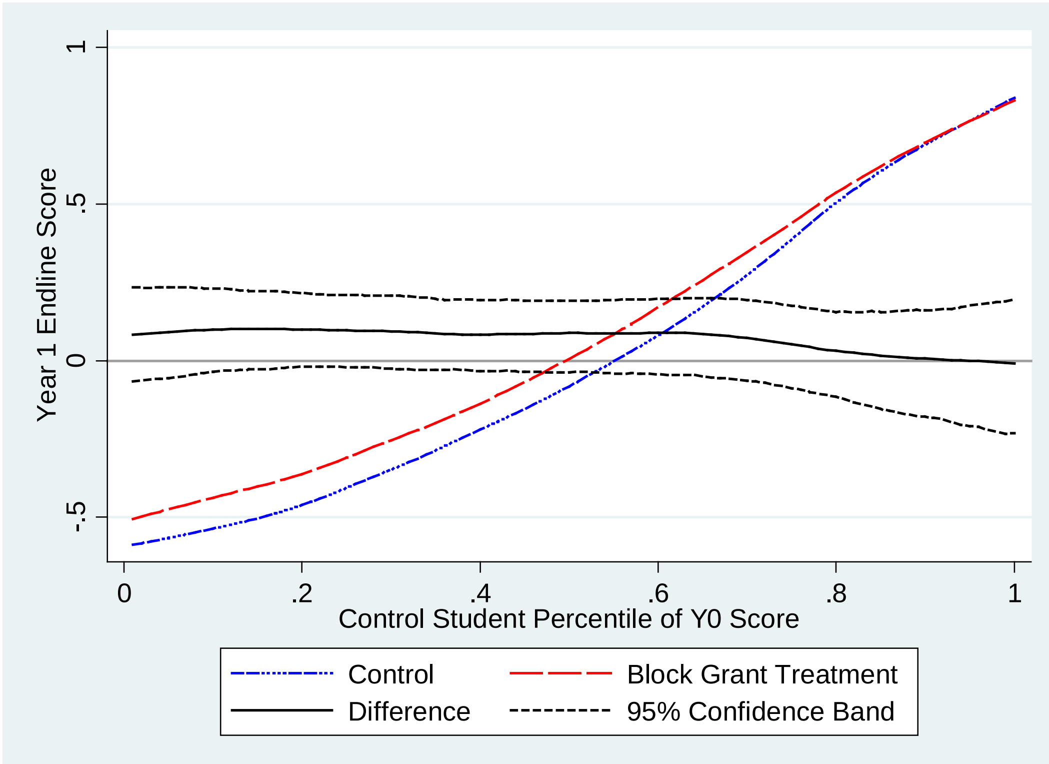

Figure 2: Average non-parametric treatment effects (Y2 on Y1)

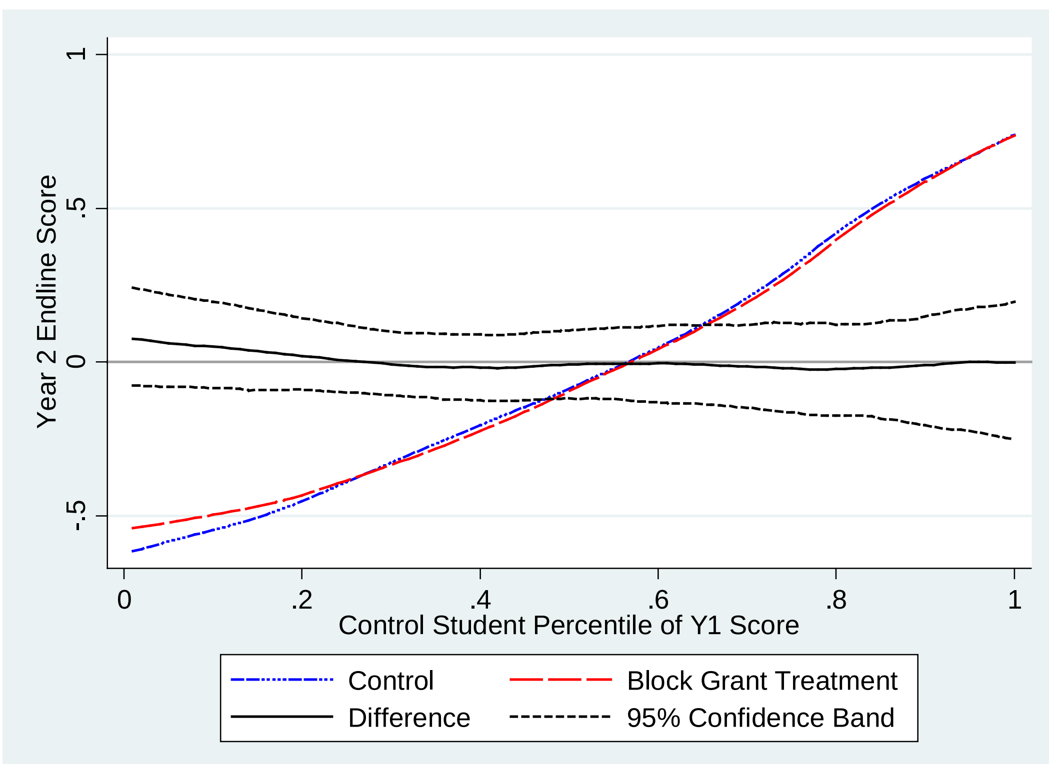

Table 1: Sample Balance Across Treatments (Andhra Pradesh School Block Grant Experiment)

Variable type	Variable	[1] Control	[2] Block Grant	[3] P-value (H0: Diff = 0)
School-level Variable	Total Enrollment (Baseline: Grades 1-5)	113.2	104.2	0.39
	Total Test-takers (Baseline: Grades 2-5)	64.9	62.3	0.64
	Number of Teachers	3.07	3.03	0.84
	Pupil-Teacher Ratio	39.5	34.6	0.17
	Infrastructure Index (0-6)	3.19	3.40	0.37
	Proximity to Facilities Index (8-24)	14.55	14.66	0.84
Baseline test performance	Math (Raw %)	18.4	16.6	0.12
	Telugu (Raw %)	35.0	33.7	0.42

Notes: The table shows the sample balance beteen the treament and control groups in the AP Block Grant Experiment.

1. The school infrastructure index sums 6 binary variables (coded from 0 - 6) indicating the existence of a brick building, a playground, a compound wall, a functioning source of water, a functional toilet, and functioning electricity.

2. The school proximity index ranges from 8-24 and sums 8 variables (each coded from 1-3) indicating proximity to a paved road, a bus stop, a public health clinic, a private health clinic, public telephone, bank, post office, and the mandal educational resource center.

3. The t-statistics for the baseline test scoresare computed by treating each student/teacher as an observation and clustering the standard errors at the school level (Grade 1 did not have a baseline test). The other t-statistics are computed treating each school as

Table 2: Spending of School Grant (Average per Block Grant School)

	Year 1		Year 2	
	Rs.	%	Rs.	%
Textbooks	110	1.1	246	2.6
Practice books	1,782	17.7	1,703	17.8
Classroom materials	2,501	24.9	2,354	24.6
Child Stationary	4,076	40.5	4,617	48.2
Child Durable Materials	864	8.6	88	0.9
Sports Goods and Others	723	7.2	577	6.0
Average Total Expenditure per Block Grant School	10,057	100	9,586	100

Notes: The table shows the average spending in Rupees and spending share in each year of the school grant.

Table 3 : Household Expenditure on Education of Children in Block Grant Schools (relative to comparison schools) over time

	Dependent variable is log of household expenditure on children's education
Block Grant School* Year 0 [β_3]	-0.023
	(0.032)
Block Grant School* Year 1 [β_4]	-0.041
	(0.027)
Block Grant School * Year 2 [β_5]	-0.212
	(0.034)***
Observations	37004
R-squared	0.135
P-value ($H_0 : \beta_4 = \beta_5$)	0.00

Notes: Household expenditure on children's education is the sum of spending on textbooks, notebooks, workbooks, pencils, slates, pocket money for school, school fees, and other educational expenses. Block Grant is a dummy denoting whether the school was a treatment school receiving the block grant or not.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4 : Impact of Block Grant on Student Test Scores (Separated by Year)

	Dependent Variable: Gain in Normalized Test Scores					
	Combined (Math & Language)		Mathematics		Language (Telugu)	
	[1]	[2]	[3]	[4]	[5]	[6]
	One-year Gain	Two-year Gain	One-year Gain	Two-year Gain	One-year Gain	Two-year Gain
Block Grant School	0.085	0.053	0.091	0.039	0.079	0.065
	[0.038]**	[0.045]	[0.042]**	[0.049]	[0.038]**	[0.046]
Observations	27704	19872	13778	9891	13926	9981
R-squared	0.269	0.325	0.293	0.325	0.254	0.238

Notes: All regressions include mandal (sub-district) fixed effects and standard errors clustered at the school level. Estimates of two-year gains do not include the cohort in grade 1 in the second year (since they only have exposure to one year of the program).

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5 : Impact of Block Grant on Student Test Scores (Pooled)

	Dependent Variable: Gain in Normalized Test Scores					
	Combined (Math & Language)		Mathematics		Language (Telugu)	
	[1]	[2]	[3]	[4]	[5]	[6]
Block Grant School [δ_1]	0.094	0.085	0.100	0.096	0.087	0.074
	(0.044)**	(0.042)**	(0.050)**	(0.048)**	(0.042)**	(0.040)*
Block Grant School * Year 2 [δ_2]	-0.058	-0.061	-0.076	-0.086	-0.039	-0.036
	(0.064)	(0.061)	(0.075)	(0.075)	(0.060)	(0.055)
Controls	No	Yes	No	Yes	No	Yes
Observations	47576	42710	23669	21244	23907	21466
R-squared	0.248	0.273	0.277	0.302	0.226	0.251
Pvalue H0: $\delta_1 + \delta_2 = 0$	0.482	0.611	0.673	0.852	0.352	0.422

Notes: All regressions include mandal (sub-district) fixed effects and standard errors clustered at the school level. Estimates of two-year gains do not include the cohort in grade 1 in the second year (since they only have exposure to one year of the program).

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6 : Impact of Block Grant on Student Test Scores (Second Year Only)

	Dependent Variable: Gain in Normalized Test Scores					
	Panel A: OLS			Panel B: Average Non-parametric Treatment Effect (Based on Figure 2)		
	[1]	[2]	[3]	[4]	[5]	[6]
	Combined	Math	Language	Combined	Math	Language
Estimated Treatment Effect	0.02	-0.008	0.047	0.01	-0.026	0.045
	[0.043]	[0.049]	[0.039]			
95% Confidence Interval	[-0.064 , 0.104]	[-0.104 , 0.089]	[-0.03 , 0.125]	[-0.105 , 0.101]	[-0.149 , 0.089]	[-0.091 , 0.164]
Observations	25772	12844	12878			
R-squared	0.246	0.302	0.206			

Notes: All regressions include mandal (sub-district) fixed effects and standard errors clustered at the school level. Confidence intervals for Panel B are constructed by calculating treatment effects using 1000 bootstrapped samples and showing the 25th and 975th largest value of these treatment effects.

Table 7: The Relative Impacts of Rule-Based Funds and the Receipt of Discretionary Funds on Test-Score Gains

	Dependent variable: Gain in normalized test-scores					
	English			Mathematics		
	[1]	[2]	[3]	[4]	[5]	[6]
Rule-Based Funds	-0.024	-0.018	-0.016	-0.010	-0.004	-0.006
	[0.031]	[0.030]	[0.031]	[0.025]	[0.026]	[0.026]
Any Discretionary Funds Received		0.103**			0.098*	
(Binary)		[0.050]			[0.048]	
Discretionary Funds			0.060**			0.029
(in Kwacha per pupil)			[0.027]			[0.024]
Square of Discretionary Funds			-0.004*			-0.001
(in Kwacha per pupil)			[0.002]			[0.002]
Observations	171	171	171	171	171	171
R-squared	0.17	0.187	0.192	0.044	0.06	0.065

Notes : The table reports the estimated effects of rule-based and discretionary funds on yearly changes in English and Mathematics test-scores. Columns 1 and 4 include only the rule-based grants, while columns 2 and 5 also include an indicator for receipt of Discretionary Funds. Columns 2 and 5 treat Discretionary Funds as a binary variable, separating schools into those who received a positive amount versus those who received nothing (rule-based funds are treated as a continuous variable). Columns 3 and 6 treat both Discretionary Funds and Rule-Based Funds as a continuous variable and includes linear and quadratic terms for the discretionary funds. All specifications include a set of geographical (province and rural/urban indicators) and school controls (including changes in school-level variables such as the head-teacher, the head of the Parent-Teacher Association and PTA fees). All standard errors are clustered at the district-level. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Are receipts of discretionary funds correlated with observable school characteristics?

	[1]	[2]	[3]	[4]
	Schools that did not receive discretionary funding	Schools that received discretionary funding	Difference	OLS Results
Total enrolment at School	887.369	989.548	-102.18	0.00005
	[677.006]	[628.406]	[118.13]	[0.0001]
Average Wealth of Students in School	-0.194	-0.046	-0.158	0.0364
	[.797]	[.735	[.138]	[0.103]
Mean Math Score at baseline	-0.019	-0.067	-0.047	-0.139
	[.443]	[.423]	[.077]	[0.0929]
Mean English Score at baseline	-0.059	-0.052	-0.007	0.105
	[.438]	[.529	[.082]	[0.0860]
Fraction Repeating	0.077	0.079	0.002	0.445
	[.065]	[.058]	[.011]	[0.597]
Fraction Dropouts in Primary	0.045	0.035	0.009	-0.555
	[.056]	[.056]	[.009]	[0.661]
DEO office <5KM	0.731	0.619	0.112	-0.066
	[.445]	[.491]	[.084]	[0.076]
PEO office <25KM	0.708	0.786	-0.078	0.129*
	[.457]	[.415]	[.075]	[0.0751]
Size of average class in school	56.295	46.908	9.38	0.151
	[38.070]	[19.240]	[6.12]	[0.166]
Observations	130	42		172
			R^2	0.04
			F-Test (All Coefficients are jointly insignificant)	1.58
			P-Value of F-test	[0.171]

Notes: The table shows the differences between schools that received any discretionary funds and those that did not. Columns (1) and (2) show the mean values and Column (3) reports the results from the mean comparisons. Column (4) reports results from a regression where we predict the receipt of any discretionary funding with school-level variables that would not have responded to the receipt of funds. The F-test cannot reject that all variables we consider are jointly insignificant, suggesting that schools that received discretionary funds were observationally similar to those that did not. For Columns (1) and (2), standard deviations are reported in brackets; for Column (3) standard errors of the difference are reported in brackets and in Column (4) we report the robust standard error after accounting for clustering at the district level.

Table 9: The Relationship between Household Spending and School Funding

	(1)	(2)	(3)	(4)
	Dependent Variable: Log of Household Spending on Child's Education			
	OLS	OLS	IV	IV
Rule Based Funds	-0.716**	-0.843***	-1.124***	-0.946**
	[0.285]	[0.252]	[0.266]	[0.460]
Discretionary Funds	0.077	0.071	0.066	0.063
	[0.109]	[0.083]	[0.091]	[0.080]
Constant	14.69***	15.52***	18.42***	16.25***
	[2.617]	[2.454]	[2.383]	[3.561]
Geographic, School, and HH Controls	N	Y	N	Y
F-stat of First Stage			23.54	10.32
Observations	1,195	1,116	1,164	1,085
R-squared	0.053	0.239	0.037	0.238

Notes: This table shows the relationship between household spending and funding received at the school, based on the sample of 34 remote schools for whom we have matched data between households and schools. We report OLS and IV coefficients for the response of household spending to rule-based and discretionary funding at the school-level. Columns 1 and 3 have do not include any controls. Columns 2 and 4 include a full set of geographical controls (province and rural dummies), household controls (child gender, age, age-squared, parental presence, parental literacy and household wealth measured through an asset index), and school controls (class-size in the school, textbooks available per child, and the number of desks and chairs per 100 children). Columns (3) and (4) are the estimated coefficients from an instrumental variable specification where we use the size of the school catchment as an instrument for per-student rule-based funding as discussed in the text. The F-statistic of the first-stage for each specification is noted. Standard errors are clustered at the school level. *** p<.01, ** p<.05.

Table 10: Household Spending and Rule-Based Allocations in the School

Funding Type		Low Rule Based Grant Schools (N=17)	High Rule Based Grant Schools (N=17)	Difference
Average Per-Child Household Expenditure (Kwacha)	Mean	17882	12022	5860***
	Standard Deviation	[26054]	[22695]	[1391]
	Observations (Households)	612	620	1232
Rule-Based funds (Kwacha)	Mean	5915	12158	-6243***
	Standard Deviation	[1733]	[2893]	[818]
	Observations (Schools)	17	17	34
Total Household and Rule-Based Funding (Kwacha)	Mean	23734	24124	-390
	Standard Deviation	[25810]	[23164]	[1396]
	Observations (Households)	612	620	1232

Notes: Rule Based-Funds show the per-student funding received under the BESSIP funding. Total Household and Rule-Based funding shows the sum of the two. The 34 schools in the sample are categorized into two equal groups with low and high rule-based funding. *** p<0.01. 1 US dollar = 3570 Kwacha on 1 September 2001.