Written Exam for the M.Sc. in Economics summer 2014

Advanced Development Economics: Micro Aspects

Final Exam

SUGGESTED ANSWERS

11 August, 2014

(3-hour closed book exam)

Please note that the exam is available only in English.

Question 1: Social Learning

Suppose that the production function takes following form:

$$q_{it} = 1 - (k_{it} - \kappa_{it})^2$$

where k_{it} is the level of input chosen by person i in period t and κ_{it} is the target level of input use. κ_{it} is not known at the time inputs are chosen. It is determined by

$$\kappa_{it} = \kappa^* + \mu_{it}$$

where μ_{it} is a normally distributed independent and identically distributed shock with mean zero and variance σ_u^2. At time t, person i does not know κ^* but has beliefs about κ^* which are distributed $N(\kappa_t^*, \sigma_{\kappa it}^2)$.

a) Define the concept of social learning and briefly describe the "Target Input" model.
b) Comment on how neighbor technology adoptions affect own adoption, and describe the two most important testable implications of this model.

The following question relates to the results in Conley and Udry (2010) "Learning about a New Technology: Pineapple in Ghana." American Economic Review, 100(1): 35-69.

c) Table 5 in Conley and Udry (2010) contains some of the main estimation results. Briefly describe the context and the key variable used for measuring the social learning. What are the main conclusions to be drawn from the results shown in Table 5? Discuss the implications of the results. [Note: There is no need to comment on the size of the effect, only the direction and the significance of associations.]

Table 5—Predicting Innovations in Input Use, Differential Effects by Source of Information
(Dependent variable: Innovation in per plant fertilizer use)

	A	B	C	D	E	F
Index of good news input levels ($M_{i,t}$)	1.05					
	(0.20)					
$M_{i,t}$ × novice farmer		1.07				
		(0.22)				
$M_{i,t}$ × veteran farmer		−0.46				
		(0.34)				
Index of good news input levels by novice farmers			−0.05			
			(0.39)			
Index of good news input levels by veteran farmers			1.05			
			(0.20)			
Index of good news input levels by farmers with same wealth				1.06		
				(0.22)		
Index of good news input levels by farmers with different wealth				−0.32		
				(0.32)		
Index of good news input levels on big farms					1.17	
					(0.19)	
Index of good news input levels on small farms					0.92	
					(0.20)	
Index of good news input levels, farmers with same soil						1.08
						(0.23)
Index of good news input levels, farmers with different soil						0.93
						(0.22)
Novice farmer		3.97	4.03	4.02	3.96	3.94
		(2.67)	(2.68)	(2.67)	(2.69)	(2.77)
Avg. dev. of geog. neighbors from previous use [$\Gamma_{i,t}$]	0.52	0.56	0.56	0.57	0.56	0.57
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Avg. dev. of financial neighbors from prev. use	0.52	0.55	0.38	0.41	0.23	0.24
	(0.59)	(0.57)	(0.58)	(0.54)	(0.62)	(0.61)
Village 1	−7.50	−8.09	−7.97	−8.10	−7.68	−7.79
	(1.22)	(1.50)	(1.42)	(1.48)	(1.39)	(1.36)
Village 2	−0.47	−1.91	−1.94	−1.98	−1.60	−1.59
	(1.53)	(2.07)	(1.99)	(2.07)	(1.99)	(2.03)
Wealth (million cedis)	0.10	0.41	0.36	0.40	0.24	0.26
	(0.25)	(0.17)	(0.18)	(0.18)	(0.21)	(0.21)
Clan 1	−2.36	−2.44	−2.43	−2.32	−2.24	−2.33
	(1.41)	(1.25)	(1.27)	(1.23)	(1.28)	(1.30)
Clan 2	−0.35	0.00	−0.10	−0.13	−0.26	−0.24
	(1.44)	(1.35)	(1.34)	(1.35)	(1.31)	(1.31)
Church 1	0.13	0.63	0.48	0.41	0.69	0.74
	(1.31)	(1.13)	(1.09)	(1.13)	(1.15)	(1.15)

Notes: OLS point estimates, spatial GMM (Conley 1999) standard errors in brackets allow for heteroskedasticity and correlation as a function of physical distance; see footnote 24 for details. Sample size = 107. A full set of round dummies included but not reported. Information neighborhoods defined using responses to: Have you ever gone to farmer ____ for advice about your farm?

Suggested answers

a) Define the concept of social learning and briefly describe the "Target Input" model.

Social learning. Techniques of production are characterized as being tacit and circumstantial sensitive. Seemingly identical techniques of production are used quite differently across producers and non-tradable inputs (land) vary in characteristics in ways that affect the performance of different technologies. When technology is tacit or circumstantial sensitive local investment in learning and innovation must take place. There are two ways of learning:

a. Learning-by-doing
b. Learning from others

The concept of social learning is given by:

"Social learning" = Learning-by-doing + Learning from others.

The target input model. A producer maximizes profit. In the model, inputs are costless, so output equals profits. The profit of a producer declines with the distance between the actual input used and the a priori unknown, optimal target level of input use. After input has been applied and output realized, the producer can deduce the target level of inputs in the given situation. But the "situation" changes (e.g., weather conditions). Hence, each round of production is an experiment which yields information about the distribution of the random target input. The model is described in detail in section II in Chapter 12 of Bardhan and Udry (1999, pp. 154-157).

b) *Comment on how neighbor technology adoptions affect own adoption, and describe the two most important testable implications of this model.*

Assume there is a "traditional" technology with riskless return q(a) and a new technology with an unknown (random) target input for which profits are increasing in the number of experiments performed by both the farmer and the farmer's neighbors. Let ι = 1 if the farmer uses the new technology and ι = 0 otherwise. The profit value function for period *t* is given by:

$$V_t(I_{t-1}, \mathrm{N}_{t-1}) = \max_{\iota \in \{0,1\}} (1-\iota_t) q_a + \iota_t E_t q_t (I_{t-1}, \mathrm{N}_{t-1}) + \delta V_{t+1}(I_t, \mathrm{N}_t)$$

In the profit value function, *I* is the cumulative number of own experiments with the new technology and *N* is the cumulative number of neighbor experiments. *E* is the expectation operator and δ is the time preference rate.

The neighbors' use of the new technology has a direct effect on the expected value of the flow of profits. The more experiments are done by the neighbors (high *N*), the higher the expected profit. Many neighbor adoptions may delay own adoption as the value of the information received from own experiments with the new technology is lower the more other farmers experiment. This can be seen by looking at the gain from the initial switch to the new technology:

$$q_a - E_0(q(0, N_0)) \leq \delta [V_1(1, N_0) - V_1(0, N_0)]$$

The LHS is the expected gain when not switching compared to switching to the new technology. The RHS is the expected increase in profits from the first own experiment. The RHS is decreasing in the number of experiments done by neighbors (N_0). If more farmers use new technology, less additional information is gained by own experiments.

The two most important testable implications of the simple model are:

1. It is possible to test directly if farmers learn from others.
2. It is possible to test whether neighbor and own experience are perfect substitutes and whether there is efficient learning.

c) *Table 5 in Conley and Udry (2010) "Learning about a New Technology: Pineapple in Ghana." American Economic Review,100(1): 35-69 contains some of the main estimation results. Briefly describe the context and the key variable used for measuring the social learning. What are the main conclusions to be drawn from the results shown in Table 5? Discuss the implications of the results. [Note: There is no need to comment on the size of the effect, only the direction and the significance of associations.]*

Conley and Udry (2010) investigate learning about a new agricultural technology among pineapple farmers in Ghana who switched to growing pineapple from traditional crops as the price of pineapple in export markets increased. A switch to growing pineapple meant that farmers were exposed to a new production technology comprising the intensive use of fertilizer and other agricultural chemicals. Conley and Udry (2010) measure the role of social learning in this process. The learning takes place as farmers learn about the optimal (target) level of inputs, such as fertilizer. To learn about the target level of inputs, farmers observe profits, input levels and growing conditions on their neighbors' farms. Farmers form expectations about the profitability of new technology depending on how close they were to the target level of input use compared to their neighbors and depending on how profitable neighbors' harvests have been with that level of input. If a farmer observes that his neighbor's profits were higher than expected at the same level of inputs, it is considered that the farmer has received good news. The opposite holds for bad news. Based on this, Conley and Udry (2010) construct the index of good news which they use to measure the impact of social learning on new technology adoption.

Table 5 shows the estimations of the relationship between the innovations in input use and the sources of information about the profitability of new technology. Column A shows that the coefficient on the index of good news input levels (M) in the farmer's information neighborhood is positive and statistically significant. The increase in M is associated with an increase in fertilizer use: Farmers tend to increase (decrease) input use when an information neighbor achieves higher than expected profits when using more (less) inputs than they previously used. Column B shows the relationship between experience and farmer's responsiveness to information on the profitability of fertilizer. There is no evidence that veteran pineapple farmers respond at all to good news about alternative levels of fertilizer use. Novice farmers, however, increase fertilizer use when M increases. Column C defines M separately for novice and veteran farmers in farmer's information neighborhood. The coefficient on M using veteran farmers' results is large and significant, and the coefficient M for novice farmers' information is not. Column D looks at the effect of social learning from neighbors in the same wealth category (both farmer and neighbor are rich or both are poor). Wealth-partitioned M is an important and significant predictor for same category neighbors but not for different category neighbors (poor learn from the poor; rich learn from the rich). Column E presents estimates split by the size of farms in i's information neighborhood. Both coefficient estimates are positive and statistically significant, suggesting that the responsiveness of input use to news from large farmers may be stronger than it is to similar news from small farmers. Finally, Column F presents estimates using the similarity/dissimilarity in the soil type (sandy or clay) between the farmer and his information neighbor. These estimates provide no significant evidence that news from others with the same soil type matters more to a farmer.

In summary, Conley and Udry (2010) show that novice farmers react to good news and they tend to react to information revealed by neighbors who are veterans and who have similar wealth. More broadly, their results show that information has value for farmers, as do the network connections through which that information flows.

Question 2: Rural Land Markets

a) Describe how a limited liability constraint, *i.e.*, a scheme in which the tenant is only liable up to his own wealth level, may affect a sharecropping contract.

The questions below refer to the analysis in Banerjee, Gertler and Ghatak (2002). "Empowerment and Efficiency: Tenancy Reform in West Bengal", *Journal of Political Economy*, 110(2), 239-280.

b) Operation Barga in India was a drive to increase tenant registration in West Bengal in India. A registered tenant could not be evicted as long as they paid their dues and the maximally legal binding due was set at 25 percent of the output. Explain and discuss the expected effects of operation Barga on agricultural productivity.
c) Describe the approach used in Banerjee, Gertler and Ghatak (2002) to test the effect of operation Barga.

Suggested answers

a) Describe how a limited liability constraint, i.e., a scheme in which the tenant is only liable up to his own wealth level, may affect a sharecropping contract.

This sharecropping model is from Chapter 6, Section III in Bardhan and Udry (1999, pp. 67-74). The model shows that binding limited liability constraints under risk neutrality gives rise to an optimization problem that is similar to the more standard sharecropping contract with risk aversion (and risk sharing). The requirement in both models is that the effort of the tenant is unobservable.

There are many tenants. Any tenant is liable up to his own wealth (w > 0) and the tenant has an outside option (m > 0). The tenant chooses effort (e ∈ [0,1]), for which there exists disutility of effort characterized by d(e), d'(e) > 0, d''(e) > 0, d(0) = 0. Output (*y*) takes a high value (*H*) with probability e and a low value (*L*) with probability 1 – e. A sharecropping contract stipulates a payment schedule as a function of the random output variable. The tenant has to pay rent to the landlord (*y*-*t*(*y*)) where *t(y)* is specified as a simple function of the two possible output levels:

$$t(y) = \begin{cases} h & \text{if } y = H \\ l & \text{if } y = H \end{cases}, \quad h > l$$

When effort is unobservable, the landlord maximizes his expected rent subject to a participation constraint (PC) (the farmer must be willing to take the contract) and an incentive compatibility constraint (ICC) (the effort must be the highest possible, given the contract):

$$\begin{aligned}
&\max_{t(\cdot)}\ e(H-h)+(1-e)(L-l)\\
&s.t.\\
&eh+(1-e)l-d(e)\geq m \quad \text{(PC)}\\
&\mathrm{e}\in \arg\max_{e} eh+(1-e)l-d(e) \quad \text{(ICC)}
\end{aligned}$$

The limited liability constraint implies that the transfer must be less than the farmer's wealth for all realizations of the output ($t(y)+w \geq 0$). Specifically, in the low output case the maximization problem is subject to a third condition (Limited Liability Condition):

$$l+w\geq 0 \qquad \text{(LLC)}$$

This says that in the bad outcome the farmer can at most pay his total wealth.

The limited liability constraint means that for poor farmers (low *w*) the landlord may not be able to set the transfer payment in the bad outcome so as to produce a sufficiently high powered contract. The result is that the gain for the farmer in the good outcome (H – h) will be lowered because the landlord will require a higher payment in the good outcome as he cannot get as much as he wants in the bad outcome. This will induce a lower effort level from the farmer and there will be allocative inefficiency for poor tenants (tenants with binding LLC) compared to better-off tenants (for whom the LLC is not binding).

b) *Operation Barga in India was a drive to increase tenant registration in West Bengal in India. A registered tenant could not be evicted as long as they paid their dues and the maximally legal binding due was set at 25 percent of the output. Explain and discuss the expected effects of operation Barga on agricultural productivity.*

Barga reduced eviction threats. There are two effects such reduced threats: Changes in bargaining power and changes in security.

The Bargaining power effect: Removal of eviction as a threat reduces the landlord's bargaining power, and forces him to offer the tenant a higher crop share, which translates into stronger incentives.

The Security effects: (i)The landlord may use the threat of eviction when output is low to induce the tenant to work harder. With Barga he cannot use the eviction threat as a discipline device which may reduce effort and efficiency. (ii) The greater security of tenure encourages the tenant to invest more since it gives him the confidence that he will stay on the land long enough to enjoy the fruits of his investment.

The overall effect of Barga on productivity is ambiguous.

c) *Describe the approach used in Banerjee, Gertler and Ghatak (2002) to test the effect of operation Barga.*

Banerjee, Gertler and Ghatak (2002) compare rice productivity (yield per hectare) in West Bengal (state in which operation Barga was implemented) and Bangladesh (Barga was not implemented) before and after the introduction of operation Barga, which was introduced in 1978. Thus, this is a basic difference-in-difference approach. Using an estimate of the fraction of rice area under sharecropping in West Bengal (about 25%), they find an increase of 51% on the productivity of registered tenants.

Question 3: Education

a) One of the UN MDGs is to reach 100% primary school gross enrolment worldwide. Discuss the potential problems (both in relation to quantity and quality of schooling) of focusing on gross enrolment rates only.

The questions below refer to the analysis and results in Angrist and Lavy (1999). “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement”, *Quarterly Journal of Economics*, 114(2), 533-573.

b) Outline the basic idea behind the identification strategy followed in Angrist and Lavy (1999). Describe Figure I and describe possible challenges with the identification strategy in their paper.

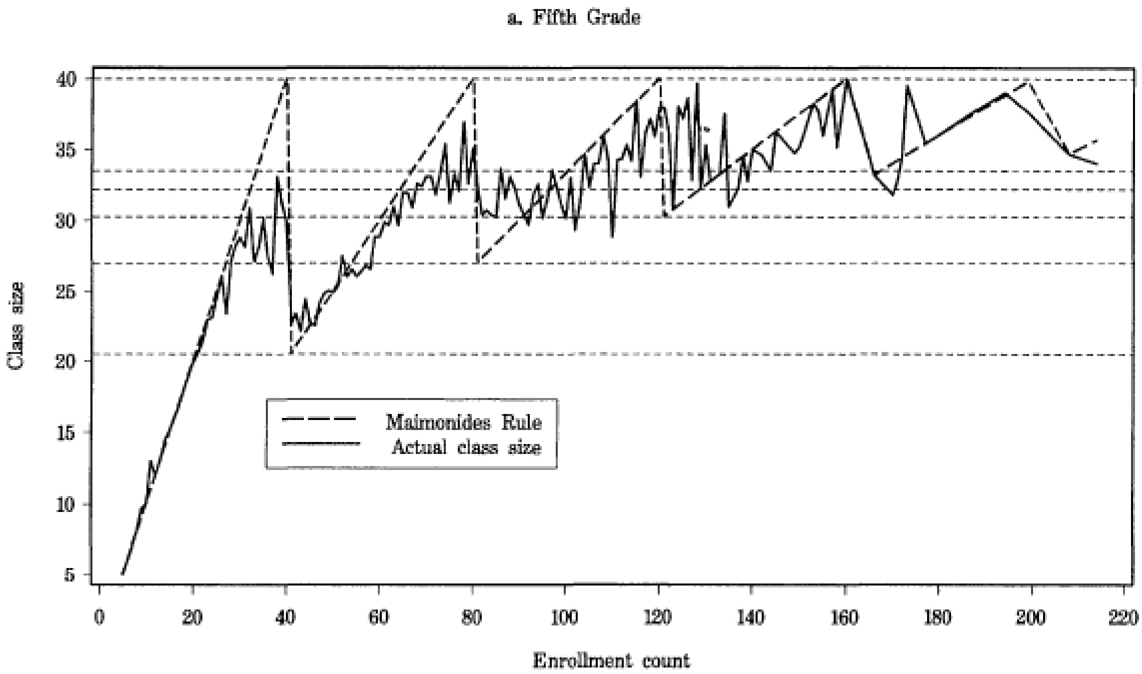

FIGURE I
Class Size in 1991 by Initial Enrollment Count, Actual Average Size and as Predicted by Maimonides' Rule

c) Table 4 contains some of the main results reported in Angrist and Lavy (1999). What are the main conclusions to be drawn from their results? Discuss the implications of the result.

TABLE IV

2SLS ESTIMATES FOR 1991 (FIFTH GRADERS)

	Reading comprehension						Math					
	Full sample				+/− 5 Discontinuity sample		Full sample				+/− 5 Discontinuity sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mean score		74.4			74.5			67.3			67.0	
(s.d.)		(7.7)			(8.2)			(9.6)			(10.2)	
Regressors												
Class size	−.158	−.275	−.260	−.186	−.410	−.582	−.013	−.230	−.261	−.202	−.185	−.443
	(.040)	(.066)	(.081)	(.104)	(.113)	(.181)	(.056)	(.092)	(.113)	(.131)	(.151)	(.236)
Percent disadvantaged	−.372	−.369	−.369		−.477	−.461	−.355	−.350	−.350		−.459	−.435
	(.014)	(.014)	(.013)		(.037)	(.037)	(.019)	(.019)	(.019)		(.049)	(.049)
Enrollment		.022	.012			.053		.041	.062			.079
		(.009)	(.026)			(.028)		(.012)	(.037)			(.036)
Enrollment squared/100			.005						−.010			
			(.011)						(.016)			
Piecewise linear trend				.136						.193		
				(.032)						(.040)		
Root MSE	6.15	6.23	6.22	7.71	6.79	7.15	8.34	8.40	8.42	9.49	8.79	9.10
N		2019		1961	471			2018		1960	471	

The unit of observation is the average score in the class. Standard errors are reported in parentheses. Standard errors were corrected for within-school correlation between classes. All estimates use f_{sc} as an instrument for class size.

Suggested answers

a) *One of the UN MDGs is to reach 100% primary school gross enrolment worldwide. Discuss the potential problems (both in relation to quantity and quality of schooling) of focusing on gross enrolment rates only.*

The student should be able to discuss the exact definition of gross enrolment rates versus the definition of net enrolment rates.

- Gross enrollment rate is defined as the number of children enrolled in a particular level of education as a percentage of the population in the age group associated with that level.
- Net enrollment rate is defined as a number of children enrolled in a particular level of schooling who are of the age associated with that level of schooling, divided by all children of the age associated with that level of schooling.

Gross enrolment rates could therefore be larger than 100% due to early or late enrolment and repetition. The answer should highlight that net enrolment rates by definition cannot exceed 100% and given that that both late enrolment and repetition rates are much higher in developing countries, the distance between gross and net enrolment rates are much higher in the developing world as compared to OECD countries. Moreover, primary school completion rates (Grade 4 survival rates) are especially low in Sub-Saharan Africa and South Asia. In relation to this issue, the good answer also discusses the problems of distinguishing between a school drop-out and a frequent absent student.

In the discussion of quality of schooling (in relation to the use of the gross enrolment rate definition), the answer could in addition consider (i) teacher absence rates, (ii) class size/pupil-teacher ratio issues (too large classes reduce the quality of schooling – increases in enrolment rates have to come hand-in-hand with investments in teachers and schools if quality of schooling is not to suffer), (iii) teacher education, experience and salary and (iv) the effect of school facilities on quality of schooling (Hanushek).

The questions below refer to the analysis and results in Angrist and Lavy (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement", *Quarterly Journal of Economics,* 114(2), 533-573.

b) *Outline the basic idea behind the identification strategy followed in Angrist and Lavy (1999). Describe Figure I and describe possible challenges with the identification strategy in their paper.*

Angrist and Lavy (1999) aim to estimate the effect of class size on educational achievement (test scores in reading and math). Since class size is not a random variable, they use the class-size function generated by Maimonides' rule to construct instrumental variables in the estimation of the class-size effect on test scores in Israeli public schools. Maimonides' rule says that a class should not be larger than 40 pupils: when 41 pupils are enrolled, the class should be split in two. This rule creates a discontinuity in the relationship between enrollment and class size at regular intervals (enrollment multiples of 40). In that way, they can match the nonlinearity or discontinuity in the relationship between the rule and the actual (observed) class size. This strategy was inspired by Campbell (1969).

The graph shows the class-size function generated by the Maimonides' rule and the actual school class sizes. At enrollment levels that are not integer multiples of 40, class size increases approximately linearly with enrollment size. But average class size drops sharply at integer multiples of 40. The figure shows that average class size almost never reaches 40 when the enrollment is less than 120, even though the class-size function predicts a class size of 40 when the enrollment is 40, 80, 120, etc. This is because schools can sometimes afford to add extra classes before reaching the maximum class size.

The problem with Angrist and Lavy's (1999) approach could be that the test scores are affected by some mechanism other than the class size. For example, both class size and instrument could be a function of the size of enrollment of cohorts. Different factors correlated with the enrollment and class size that are captured in the error term of the estimation equation are also likely to be correlated with pupil achievement. However, Angrist and Lavy (1999) assume that any other mechanism that can affect the test scores is likely to have a smoother effect – not discontinuous like the instrument. To control for any other relationship between enrollment and test scores, they include control functions of enrollment in the vector of covariates.

Angrist and Lavy's (1999) results are valid when the selective manipulation (self-selection) by parents can be ruled out. In Israel, socioeconomic status is inversely related to local population density. Better schools might face increased demand if parents selectively choose districts on the basis of school quality. At the same time, more educated parents may try to avoid overcrowded schools by moving to districts they assess will have smaller classes. This could cause correlation between unobserved parental preferences for child education and the instrumental variable used to predict the class size. The authors judge that this form of bias is small in practice. Manipulation of class size by parents is limited by the fact that Israeli pupils must attend a neighborhood school. Also, very few Israeli children are sent to private schools in order to tackle the problem of large enrollment.

c) *Table 4 contains some of the main results reported in Angrist and Lavy (1999). What are the main conclusions to be drawn from their results? Discuss the implications of the result.*

Table 4 shows the relationship between class size and test scores for reading and math. The estimation with instrumental variables shows a negative association between class size and reading achievement for fifth graders. The first column shows the estimates of a model where only the socioeconomic status of pupils is controlled for through the variable Percent disadvantaged. The estimated value of the effect of class size on the reading test scores is -0.16 with a standard error of 0.04, which is a value significantly different from zero. The same equation is estimated for the math test scores in column 7. Even though the result indicates a negative relationship between class size and math test scores, the result is not significantly different from zero. Including control variables leads to more precise estimates with the coefficient sizes similar to those for reading test scores.

The negative relationship between reading test scores and class size is confirmed in models with additional control variables, such as the enrollment and enrollment-squared. Introducing the piecewise linear trend with the slopes identical to the slope of the linear segments on the class-size function to the equation also shows the same significantly negative relationship between class size and reading test scores.

The estimation is repeated with a sample of schools in which classes deviate from the points of discontinuity by 5 pupils, giving a +5/-5 discontinuity range. The purpose of such a sample restriction is to exploit the variability in class size generated by jumps in class size at the points of discontinuity. The coefficients estimated for a discontinued sample are larger than for the full sample but less precisely estimated.

Angrist and Lavy’s (1999) results show that school resources are an important factor in improving learning outcomes.