WPS6694

Policy Research Working Paper 6694

Incentives and Teacher Effort

Further Evidence from a Developing Country

Hai-Anh H. Dang
Elizabeth M. King

The World Bank
Development Research Group
Poverty and Inequality Team
&
Human Development Network
Education Sector
November 2013

Abstract

Few would contest that teachers are a very important determinant of whether students learn in school. Yet, in the face of compelling evidence that many students are not learning what they are expected to learn, how to improve teacher performance has been the focus of much policy debate in rich and poor countries. This paper examines how incentives, both pecuniary and non-pecuniary, influence teacher effort. Using school survey data from Lao PDR, it estimates new measures of teacher effort, including the number of hours that teachers spend preparing for classes and teacher provision of private tutoring classes outside class hours. The estimation results indicate that teachers increase effort in response to non-pecuniary incentives, such as greater teacher autonomy over teaching materials, and monitoring mechanism, such as the existence of an active parent-teacher association and the ability of school principals to dismiss teachers. Methodologically, the paper provides a detailed derivation of a simultaneous ordinary least squares-probit model with school random effects that can jointly estimate teacher work hours and tutoring provision.

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

Incentives and Teacher Effort: Further Evidence from a Developing Country

Hai-Anh H. Dang and Elizabeth M. King [*]

JEL: C3, I2, O1
Keywords: teacher incentives, teacher effort, private tutoring, joint probit-OLS, random effects, Lao PDR
Sector Board: EDU, POV

[*] Dang (hdang@worldbank.org) and King (eking@worldbank.org) are economists respectively at the Poverty and Inequality Unit, Development Research Group and Education Sector, Human Development Network at the World Bank, Washington, DC. We thank our colleagues Keiko Miwa, Dominique van de Walle and Jeffrey Waite for their contributions to the design and application of the school survey in Lao PDR and for helping us interpret some of our results; and Harold Alderman, Eric Gronqvist, Yang Huang, Dean Jolliffe, Phil Keefer, Xubei Luo, Salvador Navarro, Ha Nguyen, Daniel O'Brien, Forhad Shilpi, Trang Tran, Futoshi Yamauchi, seminar participants at the North American Econometric Society Winter Meeting (San Francisco), and International Food Policy Research Institute for useful comments on earlier drafts. We also thank the National Statistical Center, Lao PDR for permission to use the data. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author and should not be attributed to the World Bank, its Executive Directors or the countries they represent.

I. Introduction

Teachers and their performance in classrooms affect the ability of any education system to produce learning results, but two problems can get in the way. One is that teachers may not be adequately prepared to teach, and, second, even when teachers are well-trained, they may not be motivated to do their best because good performance is not aptly rewarded, salaries being determined primarily by seniority, not performance, and performance may be neither monitored nor measured. To improve performance, some countries have linked at least a portion of teacher pay to criteria related to performance, usually to how well students do on tests. Opponents of this approach argue that the work of teachers is multidimensional, with only some aspects of it being measured by student test scores and that linking pay to student performance would lead to teachers teaching mainly to the test. Moreover, research suggests that non-pecuniary and implicit incentives, such as work conditions and peer pressure, may be sufficient or even more powerful in raising teacher effort.[1]

We do not assess performance-based pay in this paper, but we examine how other incentive mechanisms influence the different choices teachers make about their work. We contribute to a small, but growing, literature on the quantitative impact of incentives on teacher effort in developing countries. Specifically, we make several contributions to the literature. First, we offer new measures of teacher effort that include the number of after-school hours a teacher works, whether or not a teacher offers after-school tutoring, whether or not after-school tutoring is for pay and how many students a teacher tutors after school. Previous studies have used teacher absence rates as an indicator of effort, but except for studies that rely on unannounced school visits to measure teacher absence (Chaudhury and others 2006), absence records are generally thought to be unreliable because of the strong incentive for teachers and school principals to

[1]See Podgursky and Springer (2007) and Neal (2011) for recent reviews of teacher performance pay programs.

underreport absences to a central education agency. In contrast, work hours outside of the classroom to prepare lessons and provision of after-school tutoring are not mandated, so we argue that the incentive to misreport them is likely to be weaker, at least in the Lao context.

Furthermore, if teacher effort were regarded as a continuous and unobserved (latent) index variable that is manifested in terms of both absence and number of work hours (or tutoring activities), teacher work hours would approximate this variable better since work hours are a continuous variable with multiple values while absence is just a binary variable with two values.[2] The number of work hours or different aspects of tutoring activities as a measure of teacher effort is especially relevant in places where absenteeism is not excessively high. Incidentally, our new measures of teacher after-school tutoring activities make this paper a useful contribution to another emerging literature on private tutoring.[3] Most existing studies on after-school tutoring, however, focus on student demand and use of such services, and we add a little-explored angle by investigating instead the supply side of this phenomenon as part of a teacher's work choice.

Second, we use reported delays in the payment of teacher salaries to capture pecuniary incentives,[4] and the extent of teacher autonomy, parent- teacher associations' (PTA) activities, and a school principal's authority to measure non-pecuniary incentives. While these incentives represent the current practice in the education system in Lao PDR and are relatively amenable to policy influence in this country as well in other developing countries' context, most of them have

[2] Models with unobserved latent index belong to a general group of models called limited dependent variable models. See, for example, Maddala (1983) for a general treatment of such models in econometrics.
[3] Private tutoring (or supplementary education) is widespread and can be found in countries as economically diverse as Cambodia, the Arab Republic of Egypt, Japan, Kenya, Morocco, Romania, Singapore, Turkey, the United States, and the United Kingdom. A survey of the prevalence of tutoring in 22 developed and developing countries finds that in most of these countries, 25–90 percent of students at various levels of education receive private tutoring, and private tutoring can have a positive impact on student performance (Dang and Rogers, 2008). In several UNESCO and ADB publications, Bray (2009) and Bray and Lykins (2012) argue that the phenomenon of private tutoring deserves far more attention from both policy makers and researchers than it has been given.
[4] Salary delays may also be a non-pecuniary incentive if teachers interpret it as, say, lack of appreciation for their job; however, given the common occurrence of this phenomenon in Lao PDR alongside generally high teacher satisfaction, it is much more likely for salary delay to function as a pecuniary incentive.

not been used in previous studies. We examine how teachers change their level and type of effort in response to these different incentives.

Third, we use a simultaneous OLS-probit econometric model to jointly estimate how pecuniary and non-pecuniary incentives affect two broad measures of teacher effort—after-school work hours and after-school tutoring activities—thus obtaining more efficient estimates than a single-equation estimation approach.[5] We incorporate a school random effects component in this model, which can help capture heterogeneity at the school level as well reduce the bias caused by unobservable school characteristics. This type of model appears to be infrequently used in the economic literature, thus we provide a detailed derivation of its likelihood function and first derivatives that can be applied to other similar econometric issues.

Finally, to our knowledge, scanty quantitative evidence currently exists on the education system in Lao PDR, one of the poorest countries in Asia. Thus our analysis using a nationally representative school survey database is relevant to policy making in this country as well as adds new evidence to the current global stock of knowledge.

We find that both pecuniary and non-pecuniary incentives urge teachers to exert more effort in after-school work hours or after-school tutoring activities or both. In particular, one month of salary delay reduces the odds of teachers providing tutoring lesson free of charge by 40 percent and the number of their tutees by 30 percent. The existence of a PTA increases the probability that teachers offer tutoring by 0.12 percent, while the freedom to develop teaching materials induces teachers to work three additional hours per week. Incentives such as teacher autonomy over the teaching method or the principal's power to dismiss teachers lead teachers to substitute

[5] Joint models for discrete and continuous outcomes are, nonetheless, widely used in the statistic literature (see, for example, Gueorguieva and Agresti (2001) and Liu, Daniels and Marcus (2009)). The closest version to this model without the random effects in the economic literature is perhaps the instrumental probit model (see, for example, Wooldridge (2010)).

effort away from the less visible activity of preparing for classes and towards the more visible activity of after-school tutoring.

This paper consists of eight sections. In the next section, we briefly review the related literatures on teacher incentives and effort with a focus on developing countries. Section III, which provides the analytical framework, is followed by a description of teacher incentives in Lao PDR and our data in Section IV, and Section V discusses estimation results. Sections VI and VII provide further investigation of teacher provision of tutoring and the effect of PTAs on teacher effort. Section VIII summarizes our findings and discusses policy implications.

II. Literature Review

While ample evidence exists on teacher incentives in richer countries,[6] much less is known about teacher incentives in developing countries. It is tempting simply to apply the findings in richer countries to poorer countries, but there are important differences between rich countries and developing countries, including shortage of trained teachers, poor teacher quality and a weak monitoring system (Glewwe and Kremer, 2006). In this section, we briefly review the most relevant studies on teacher incentives in developing countries.

Teacher effort is difficult to measure in part because it happens behind classroom doors, away from the eyes of school inspectors and even school principals, and in part because effort has several dimensions. Absenteeism has been used as a proxy measure of effort in previous studies, but it is typically an unreliable indicator because it runs into problems associated with self-reporting and possible collusion among teachers and other school actors. To overcome this measurement problem, a survey of schools conducted in Bangladesh, Ecuador, India, Indonesia, Peru, and Uganda used *unannounced* school visits and checks of teachers present against a roster

[6] See Podgursky (2011) and Neal (2011) for recent reviews.

of teachers who were supposed to be in classrooms (Chaudhury et al., 2006). The survey revealed high teacher absence rates—from 11 percent to 27 percent of teachers were missing from the classroom without an official excuse. Analyses of the data indicate that teacher position and birthplace (i.e., teachers being born in the same districts as their school) are negatively associated with absence, as are school and community characteristics such as poor school infrastructure, school distance from the government education office, and parental literacy rates (Chaudhury et al., 2006). The authors conclude that weak monitoring exacerbates teacher absence, especially when excessive absenteeism has no negative consequence for offenders.

Various incentives to improve teacher effort and performance have been evaluated, but with mixed results. In a randomized study in India, Duflo, Hanna and Ryan (2012) find that the use of financial incentives and careful monitoring (through daily photos of the teacher with her students) reduce teacher absence rates by 21 percentage points relative to a control group. In another randomized study, Kremer, Miguel and Thornton (2009) find that greater parental monitoring increases teacher attendance in Kenya.[7] However, these results differ from those of other studies on the same countries which conclude that teacher effort does not change despite greater participation by beneficiaries (e.g. students and parents) in monitoring in India (Banerjee et al., 2010), or performance-based financial incentives in Kenya (Glewwe, Ilias and Kremer, 2010).[8]

[7] Analyzing a survey on public health workers in Lao PDR, Yamada, Sawada and Luo (2013) also find that timely payment of wages and efficient monitoring of workers' attendance can reduce absenteeism, even though absenteeism rate among these workers is low compared to other developing countries.

[8] It is interesting to see these mixed results for the same countries. While there are perhaps a large number of randomization studies for these countries, Banerjee et al. (2010) also observe that the implementers of these incentives matter: incentives implemented by nongovernmental organizations (NGO) were effective while those implemented by government bureaucrats were not. See also Glewwe, Holla and Kremer (2009) for a review of teacher incentives, and Kremer, Brannen and Glennerster (2013) for an overview of education challenges in developing countries.

Studies using teacher work hours are few because of lack of reliable data, and the few existing studies similarly offer no conclusive evidence. An evaluation of the Education with Community Participation Program (EDUCO) in El Salvador estimates the impact of delegating to the local community the tasks of hiring and firing teachers and of day-to-day monitoring, and finds that teachers in EDUCO schools or classrooms taught up to 9.7 hours more per week than their counterparts in non-EDUCO schools (Sawada and Ragatz, 2005). In contrast, another evaluation of a similar community school program in Honduras (PROHECO) finds no significant impact on teacher work hours (Di Gropello and Marshal, 2005).[9]

Our paper is related also to a nascent literature on private tutoring.[10] The research on private tutoring mainly examines the *demand* side of tutoring, which includes the factors that determine student attendance at private tutoring classes (see Dang and Rogers (2008) for a review), rather than the *supply* side of tutoring. To our knowledge, the only study that offers quantitative evidence on this topic is by Jayachandran (2013) who finds that after-school tutoring in Nepal has distortional effects on classroom hours, such as teachers reducing their in-class teaching load and teaching time.[11] That paper, however, assumes as given the decision to provide tutoring at the school level while our paper examines why individual teachers provide private tutoring in the first place.

---

[9] Other papers that look at teacher work hours as a measure of teacher effort in high-income countries include Waterreus and Dobbelsteen (2001) and Lavy (2009). Using an instrumental variable model, Waterreus and Dobbelsteen find that wages have a statistically significant impact on Dutch teacher working hours with the uncompensated wage elasticity being 0.2 for males and 0.4 for females. And using a randomized experiment, Lavy finds evidence that performance pay incentive induces teachers to change teaching methods and work longer hours.

[10] Private tutoring has been found to enhance various measures of student academic performance including student test scores in India (Banerjee et al., 2010), Israel (Lavy and Schlosser, 2005), the United States (Jacob and Lefgren, 2004), and grade point averages (GPA) ranking in Vietnam (Dang, 2007) as well as the quality of universities students attend in Japan (Ono, 2007). But see also Dang and Rogers (2008) for a review of other studies that do not find statistically significant impacts of private tutoring on student performance.

[11] Some theoretical evidence (Biswal, 1999) and qualitative evidence (Silova and Bray 2006) also point to the corrupt behaviors by teachers who reduce their in-class materials and force their students to attend their tutoring classes to make up for this loss in instruction.

III. Analytical Framework and Empirical Model

III.1. Analytical Framework

We start with the familiar model of labor supply choice between work and leisure in which a certain number of work hours is chosen to sustain a desired level of consumption. As is the case with most workers, however, teacher effort is determined not only by wages but also by nonwage aspects of one's work, such as the quality of work conditions, prestige and recognition which give teachers a certain amount of satisfaction or pleasure. Given this, we express teacher effort (E) as the function

$$E= f(P, A; Z) \tag{1a}$$

where P is a vector of pecuniary incentives (e.g., wages), A is non-pecuniary incentives (e.g., greater autonomy, monitoring), and Z is a set of characteristics of the teacher and the school. Pecuniary incentives are desired because they allow increased consumption levels; non-pecuniary incentives such as professional recognition and greater teacher autonomy lead to higher satisfaction.[12] Whether increasing P or A, or both, leads to significantly greater teacher effort depends on whether teachers value these incentives more than leisure, and policymakers can use either P or A, or both, to increase teacher effort. A singular focus on performance pay is thus suboptimal, but identifying the most effective mix of P and A is an empirical issue.

In Lao PDR, as in many countries, the salary scale of public school teachers is given and known, and increases in pay are often independent of individual performance level. But even

[12] Both theoretical and empirical evidence points to the role of non-pecuniary incentives in increasing teacher effort. For example, Dixit (2002) offers a principal-agent theoretical framework to interpret teacher incentives, and he considers teachers as "motivated agents…[that]…enter the teaching profession for idealistic reasons or because they enjoy working with children". Banerjee, King, Orazem and Paterno (2012) find evidence that in Pakistan teacher and student attendance are mutually reinforcing, controlling for the endogeneity of these behaviors; that is, that the most powerful factor that raises teacher attendance is the attendance of the pupils, and the most important factor influencing pupil attendance is the presence of the teacher because these actors together produce the shared goods of teaching and learning. In rural areas where the majority of teachers live in the same village where they are teaching, this perspective on the behaviors of teachers and pupils may be particularly salient.

where some pecuniary reward is available for good performance, such a reward is likely to be given on the basis only of effort that can be observed by the principal (or by students' parents), such as perfect attendance and unpaid after-school tutoring hours. Effort that is generally unobservable (e.g., hours for after-school class preparation) is not likely to be rewarded directly but, if manifested through high student performance and a reputation for being a good teacher, may be rewarded ultimately. We next operationalize this framework for our empirical model.

III.2. Empirical Model

Assuming that classroom hours are prescribed across the school system and enforced by a central agency (and that absences are either too difficult to ascertain or not excessively high), we posit that teacher effort consists of the time spent on after-school preparation and tutoring. Thus, for teacher *j*, *j*=1,…, J, at school *i*, *i*=1,…, I, let H_{ij} and Y_{ij}^* be continuous variables indicating after-school work hours and after-school tutoring hours, respectively; we then express teacher effort as

$$E_{ij} = H_{ij} + Y_{ij}^* \tag{1b}$$

However, while we have reported after-school work hours, we do not have data on the number of hours for after-school tutoring; we know only whether or not a teacher provides tutoring after school and whether or not she does so for pay. Thus following previous studies on time use (e.g., Glick (1999), Kimmel and Connelly (2007)),[13] we jointly estimate H_{ij} and Y_{ij}^* in a simultaneous equation framework. However, we do not have a direct measure of Y_{ij}^* and only observe whether teachers provide tutoring. Thus let T_{ij} be a binary variable that equals 1 if the

[13] See also Aguiar et al., (2012) for a recent review of the related literature.

teacher gives after-school tutoring and 0 otherwise, we then estimate the reduced-form version of equation (1b)

$$
\begin{aligned}
H_{ij} &= \alpha_P P_{ij} + \alpha_A A_{ij} + \alpha_Z Z_{ij} + c_i + \upsilon_{ij} \\
T_{ij} &= 1[Y_{ij}^{*} \geq 0 \mid P_{ij}, A_{ij}, Z_{ij}] = 1[\beta_P P_{ij} + \beta_A A_{ij} + \beta_Z Z_{ij} + c_i + \varepsilon_{ij} \geq 0]
\end{aligned} \tag{2}
$$

where 1[.] is the indicator function. Since after-school work hours (H_{ij}) is a continuous variable and tutoring provision (T_{ij}) is a binary variable, we estimate (2) using a simultaneous OLS-probit model.[14]

To simplify our notation, we represent P_{ij}, A_{ij}, and Z_{ij} together as a vector of the observables X_{ij}. Teachers in the same school are likely to share unobservable but correlated characteristics (e.g., having similar income ranges or characteristics), thus we control for this within-school correlation by including a school random-effects component (c_i) in the model. Conditional on X_{ij}, c_i is assumed to have a normal distribution with mean zero and variance σ_c^2.

We assume that conditional on X_{ij} and c_i, the error terms ε_{ij} and υ_{ij} are independent across teachers, but are normally distributed with zero mean vector and covariance matrix $\sum=$ $\begin{bmatrix} 1 & \rho\sigma_V \\ \rho\sigma_V & \sigma_V^2 \end{bmatrix}$ for each teacher. The variance of ε_{ij} is normalized to 1, the variance of υ_{ij} is σ_V^2,

[14] While we have data on three types of teachers (i.e. those who tutor for pay, those who tutor without pay and those who do not tutor), we aggregate the data into two types of teachers, where teachers tutor or do not tutor for three reasons. First, we can avoid further data measurement issues since tutoring for pay can be sensitive information and teachers may underreport when they tutor for pay; second, teachers mostly tutor without pay (62 percent, Table 1); and finally, this makes the modeling of the joint equation framework more tractable. In addition, it may not be easy empirically to get the model to converge since less than 5 percent of the teachers tutor for pay. However, we will come back to this issue when looking at the determinants of teacher tutoring with or without pay in a later section.

It is also possible that teachers may decide whether to provide tutoring lessons first before deciding on how many after-class hours to work, or vice versa. However, this approach would involve estimating a nested model with the probit model (for the tutoring decision) at the first stage and the linear model (for the number of hours to work) at the second stage. This model would even be more complicated if the school random effects is to be incorporated, as we do in our model. Furthermore, given the available data it is not easy to come up with good instruments for identification of this nested model. Thus we use a simultaneous equation framework instead and leave this approach for further research. See also Dang (2007, 2008) for an application of the related joint Tobit-Ordered probit model.

and the intra-class correlation for the same teacher is represented by ρ. If ρ is statistically significantly different from 0, then the two equations in (2) should be estimated using a simultaneous equation framework to achieve the most efficient estimates. The stronger ρ is, the more efficiency is gained by this simultaneous equation framework; otherwise, these two equations can be estimated separately. In this particular context, ρ measures the strength and the direction of the relationship between the unobserved teacher characteristics that affect teacher decisions to provide tutoring and those that influence after-school work hours. A negative value for ρ would mean a tradeoff between the two activities, given the total time constraint.

Given the above assumptions, the conditional log likelihood function for the entire sample can be written as (see Appendix 1 for more details),

$$\ln L\left(T_{ij}, H_{ij} \mid X_{ij}, c_i; \alpha, \beta, \gamma, \sigma_{\upsilon}, \sigma_c, \rho\right) = \sum_{i=1}^{I} \log\left(\prod_{j=1}^{J} f(T_{ij}, H_{ij} \mid X_{ij}, c_i) \right)$$

$$= \sum_{i=1}^{I} \log\left(\int_{-\infty}^{\infty} \left[\prod_{j=1}^{J} f(T_{ij}, H_{ij} \mid X_{ij}, c) \right] \frac{1}{\sigma_c} \phi\left(\frac{c}{\sigma_c} \right) dc \right) \qquad (3)$$

where $f(.)$ and $\phi(.)$, respectively, stand for the joint density and the standard normal density functions.

Before turning to estimation, we note here that while we observe the impact of incentives on after-school work hours, we do not have a direct measure of the number of hours for after-school tutoring. If incentives exert a positive effect on both components of work (or exert a positive effect on one component and have no effect on the other), then clearly total work hours (effort) would increase. However, if incentives exert a positive effect on one component and a negative effect on the other components (i.e., incentives induce a substitution away from one activity towards another), it is unclear whether total effort would increase or decrease. In such cases, total effort could increase if the increase in hours spent in one type of work exceeds the reduction in the other, and decrease otherwise. Still, since single-equation estimates are asymptotically consistent but inefficient (Wooldridge, 2010, p.188), if the estimated coefficients on these incentives based on the single-equation method are statistically significant in one regression (say, tutoring) in the same direction as those in the simultaneous equation framework but insignificant in the other (say, work hours), these conservative estimates obtained from single-equation estimation method could provide some hint on the direction of change for total effort.

The integral in (3) is amenable to maximization by maximum simulated likelihood (MSL) estimation.[15] Appendix 1 provides the derivations for the likelihood function as well as the first-order derivatives that can be used to maximize the likelihood function. We write our own program to maximize this likelihood function in Stata (see, for example, Gould, Pitblado and Poi, 2010) using Halton draws with antithetics.[16] We use Stata's maximization method d1 and the combined Berndt-Hall-Hall-Hausman and Broyden-Fletcher-Goldfarb-Shanno algorithms. We discuss the choice of variables in the next section.

IV. Data Description and Background on Teacher Incentives and Effort

IV.1. Data Description

Our analysis draws primarily on data from a school survey designed by one of us;[17] it was fielded in conjunction with the Laos Expenditure and Consumption Survey 2002-2003 (LECS3)

---

[15] See, for example, Train (2009) for a textbook treatment of maximum simulated likelihood estimation. An alternative maximization method is Gauss-Hermite quadrature (Butler and Moffitt, 1982), which can reduce computer time but results can be more unstable. We tried Gauss-Hermite quadrature in a previous draft and find qualitatively similar results. A Stata program that estimates the likelihood function in (3) is available upon request.
[16] We obtain the Halton draws with antithetics using the user-written routine "mdraws" (Cappellari and Jenkins, 2006). Halton draws with antithetics are found to increase performance of the estimation results for multivariate probit models compared to Halton draws without antithetics (Sandor and Andras, 2004).
[17] One of the authors, Elizabeth King, designed this survey with Dominique van de Walle. Keiko Miwa (World Bank) and staff in the Ministry of Education, Lao PDR, were also involved in the final design and pilot-testing of the survey.

and used the same sampling frame. LECS3 is a nationally representative survey that contains a household questionnaire and a village questionnaire. The school survey collected detailed information from each primary school serving the catchment area for the households surveyed in LECS3. If a village did not have a school at the time of the survey, the closest school attended by most children in that village, usually located in a neighboring village, was covered by the survey. Information about the school principal or head and data on the school were collected through interviews of the school principal. Individual teachers were interviewed also and were asked whether they offered tutoring lessons to students outside of school hours, and if they did, whether or not those tutoring lessons were for pay and the number of students they tutored. Teachers were also asked about the number of hours they spent preparing teaching materials and grading homework. [18] After deleting the observations with many missing values, our analysis file covers more than 1,500 teachers in 322 schools. To test whether this final sample resembles the population of schools, we compared summary statistics of similar variables from our estimation sample with those from a national school census data for the same year. For example, the female teacher ratio and the student-teacher ratio are, respectively, 0.52 and 30.4 in our final sample, which are almost identical to 0.51 and 31 from the 2002-2003 school census database (ESITC, 2011).

The school survey collected information related to pecuniary and non-pecuniary incentives. Our measure of pecuniary incentives (P) is the number of months a teacher's salary was in arrears in the past year; again, we focus on salary delay rather than salary levels because the salary scale is given. Non-pecuniary incentives (A) consists of two types of measures: The first set includes binary variables reflecting the level of teacher autonomy in a school: whether

[18] Although the survey did try to measure teacher absence rates, there were more missing data and the rest hardly had any variation.

teachers are allowed to choose their teaching method, develop their teaching materials, adapt the school curriculum to local conditions, and set the standards for their student promotion. The second set consists of monitoring mechanisms in the school: the principal's authority as measured by whether the principal can dismiss teachers, set teachers' school hours and evaluate teacher performance, as well as the school's distance to the district education bureau and the existence of a school PTA. The control variables (Z) reflect characteristics pertaining to the individual teacher and to the school and can be roughly categorized into teacher personal characteristics, school-specific teacher characteristics, and other school characteristics.[19]

One drawback of our data is that teachers' weekly after-school work hours and tutoring activities are self-reported. It is quite possible that after-school work hours, which tend to be invisible to the principal and other monitors, are over-reported, and after-school tutoring for pay may be under-reported, but because there are no official rules about these activities, as opposed to classroom hours, we assume that there is more signal than noise in these self-reported data.

IV.2. Overview of Teacher Incentives and Effort

Lao PDR has witnessed steady economic growth rates averaging seven percent in recent years, but it remains one of the poorest countries in East Asia with a per capita income level of US$1,399 in 2012 (World Bank, 2013). The country is ethnically diverse with more than 50 ethnic groups, with the Lao-Tai being the largest ethnic group, accounting for 58 percent of the

---

[19] Teacher characteristics include teacher's gender, ethnicity, educational level, and teaching experience. School-specific teacher characteristics include the number of grades and students allocated to the teacher, and whether the teacher has a guidebook, whether the teacher was born in the same district that her school is located, or whether the teacher is a civil servant (i.e. have permanent teaching position). School characteristics include distances from the school to the provincial capital, the nearest paved road, whether the school offers multi-grade teaching, and a school fee index. We construct this school fee index by aggregating six different school fees (including for tuition, sports, examination, book rental, and others), with each fee getting a score out of three depending on whether or not students are obliged to pay: 0 for no fee, 1 for optional fee, and 2 for compulsory fee; thus the higher the index, the more fees students have to pay. We did not control for teacher salary because this variable is missing for a number of observations. But note that teacher salaries in Lao PDR are tightly compressed and can be mostly represented by their education levels and teaching experience.

total population. There are about 31,000 public primary school teachers in Lao PDR, of which 51 percent are female and 77 percent are Lao-Tai, serving approximately 909,000 primary students (ESITC, 2011). Only 6.7 percent of the teachers in our sample were absent on the day of the interview and principals could explain the reasons for most (91 percent) of the absences.[20] Thirteen percent of teachers offer tutoring lessons to their students; eight percent do not charge any fee (Table 1). Teachers generally spend an average of 13 hours per week preparing for class and grading homework. Although Figure 1 shows that the distribution of teachers' after-school work hours (solid line) has somewhat a long right tail or positive skewness compared to the normal distribution (dashed line), over-reporting of work hours does not appear to be a serious issue. More than half of all teachers report putting in at least 10 after-school work hours per week, and 6 percent report zero after-school work hours. We will return to the question of over-reporting in a later section with robustness checks on estimation results.

[Table 1 about here]

[Figure 1 about here]

Teachers report frequent delays in salary payments.[21] They report a delay of 1.8 months, on average, with a maximum of up to 7 months; 64 percent report a delay of one month or more.[22] Interestingly, there seems to be a relationship between these salary delays in a school and average number of students tutored after-school by teachers in those schools (Figure 2). For

[20] A recent survey found that nineteen out of 20 teachers interviewed in a recent survey said they were satisfied with their overall work, and absenteeism was estimated at less than 9 percent (World Bank, 2008), a number which is much lower than in other countries (Chaudhury *et al*., 2006).

[21] Indeed, delays in salary payment are serious enough that teachers have raised the issue through the hotline to a recent National Assembly sitting (Vaenkeo, 2010).

[22] Authors' calculations from the school survey. These numbers are generally consistent with those from another recent Public Expenditure and Tracking survey implemented by the World Bank (World Bank, 2008; Benveniste, Marshall, and Santibañez, 2008). Also note that teacher salaries in Lao PDR are low and tightly compressed. The average monthly salary for primary teachers, including the base salary, bonuses, and family allowances is about 390,000 Kip (US$39) and is based on such factors as educational qualification, duty location and assignment, and teaching experience. For example, teachers teaching in remote and mountainous areas can get supplements up to 20 percent of their net salaries, and teachers teaching multi-grade classes can get supplements up to 50 percent of their base salaries (Benveniste, Marshall, and Santibañez, 2008).

example, teachers tutor 1.7 students, on average, in schools that do not experience salary delays, only 0.9 students in schools that experience a 2-month delay, and 0.03 students where there is a 4-month- or-more delay.

[Figure 2 about here]

Overall, teachers claim to have several non-pecuniary incentives—autonomy to develop teaching materials (41 percent), choose teaching methods (31 percent), set standards for student promotion (27 percent), and adapt the curriculum to local conditions (14 percent). Few principals have the power to dismiss teachers (9 percent), but some can set teachers' work hours (24 percent) and most can evaluate teacher performance (79 percent). Most primary schools (87 percent) have a PTA.

V. Impacts of Pecuniary and Non-Pecuniary Incentives on Teacher Effort

V.1. Estimation Results

Estimation results of the impact of different types of incentives on private tutoring and after-school work hours are obtained from a joint OLS-probit model with school random effects (Table 2). For a robustness check of estimation results and for comparison purposes, three models are built sequentially with blocks of variables. The most basic model (Model 1) contains the incentive variables, plus controls for teacher personal characteristics and whether the school is located in an urban or rural area. Model 2 adds school-specific teacher characteristics, and Model 3 adds other school characteristics. Model 3 is the most complete model and is our preferred model. Table 2 provides the coefficient estimates from the probit regression for the tutoring equation and the marginal effects derived from those estimates.[23] For the after-work

[23] The marginal effects are obtained by averaging the predicted probabilities over all observations, and we do not evaluate the marginal effects at the means of the variables (where c= 0) since the marginal effects obtained in this way may only capture a small fraction of the population (Wooldridge, 2010) and thus are not representative of the

hours equation, the coefficient estimates are reported which can be readily interpreted as in an OLS regression.

[Table 2 about here]

The intra-class correlation term ρ (ρ_teacher) for teachers is highly statistically significant (1 percent level) in Models 1 to 3 and ranges in large values between -0.6 and -0.7, suggesting that we indeed should use a joint model for analysis. A negative value for ρ, as discussed above, indicates that unobserved teacher characteristics have opposite effects on tutoring and after-school work hours, that is, teachers who tutor work fewer after-school hours, and vice versa.

Estimation results are consistent across the models. Delays in receiving salaries, our measure of pecuniary incentives, have the expected negative impact on teachers providing tutoring and after-school work hours, but the results are not statistically significant. Controlling for teacher and school characteristics, salary delays do not seem to deter teacher effort. However, higher levels of some non-pecuniary incentives—teachers being able to choose their teaching method and having a school PTA—increase the probability that teachers offer tutoring lessons by between 11 and 34 percent. Allowing teachers to develop their own teaching materials induces them to work approximately 3 hours more per week preparing for classes, presumably time spent developing the materials. A longer distance to the district education bureau is associated with more, not fewer, teacher work hours, but the coefficient is small and not statistically significant when more robustness checks are done.

Consistent with our theoretical discussion, the impact of some non-pecuniary incentives on total teacher effort—that is, on both tutoring and after-school work hours—is not clear cut.

data. In addition, generally for probit models, since the means of the variables may not correspond to any observed values in the population, and the range of an independent variable may correspond to the region of the probability curve that is nonlinear, averaging over observations can be the preferred method. See, for example, Long (1997) for more discussion on the marginal effects in the probit model.

Controlling for other factors, while being able to choose their teaching method is associated with an 11-percent higher likelihood of tutoring and 4 fewer work hours preparing for classes. Where the school principal has the power to dismiss a teacher, teachers are 34 percent more likely to offer after-school tutoring but work 2 fewer hours per week for class preparation.

A couple of remarks are in order here. First, after-school tutoring is a more observable activity that could have a better chance of impressing the school principal than more hours spent preparing teaching materials at home. Teacher provision of tutoring classes is a very noticeable activity in villages in Lao PDR, which are tightly knit communities with a strong communal life.[24] Indeed, where the principal can evaluate teacher performance, teachers work two fewer hours but are more likely to provide tutoring even though the latter result is not statistically significant (i.e., the coefficient in the tutoring regression is also positive but not statistically significant).[25] Second, since we do not have data on the number of hours teachers spend on after-school tutoring, we cannot precisely measure the degree to which this strategic effort substitution happens. However, as discussed in a previous section, some evidence based on estimation results with single-equation method can help indicate the direction of total teacher effort.[26]

[24] See, for example, Ireson (1996) and Rigg (2009) for more discussion on village identity and culture in Laos.
[25] This result is consistent with previous studies for the United States, where school principals can evaluate teacher performance, especially the worst- and best-performing ones (Jacob and Lefgren, 2008), and that teachers do improve their long-run performance with teacher evaluation (Taylor and Tyler, 2012).
[26] We also control for other characteristics as indicated in (3). For example, female teachers are 2 percent less likely to provide after-school tutoring, perhaps because tutoring which is usually held in school competes with family time for female teachers. Teachers with an upper secondary or vocational or college degree (compared to teachers with a lower secondary degree or less) are between 3 and 5 percent more likely to tutor students. Teachers with a vocational degree or in a permanent civil position work on average one or two more hours per week, but these are marginally statistically significant at the 10 percent level. Teachers with more students in their classes are more likely to offer tutoring and work longer hours after school, but the effects are modest –one additional student in class leading to just 0.1 percentage point increase in the probability of tutoring or 0.1 hours more per week. It is particularly interesting that teachers given a guidebook work four more hours per week, suggesting that teachers increase their effort when they have the necessary teaching tools. In addition, teachers in rural schools appear to substitute after-school working hours for tutoring, although the urban-rural dummy variable loses its significance in Model 3, indicating perhaps that other school location characteristics such as distances to the provincial capital and the nearest paved road are better measures of the ability to monitor teachers.

V.2. Robustness Checks and Further Analysis

Single-equation estimation method

As discussed above, the simultaneous OLS-probit model is much more efficient than the single-equation estimation method since the intra-class correlation term ρ for teachers is highly statistically significant and has large magnitudes. Single-equation estimation method can provide asymptotically consistent estimation results assuming our estimation sample is large enough (under the framework postulated with equation (2)). Thus as a robustness check, we re-estimate the tutoring and work-hours equations using single-equation probit and OLS regressions, respectively, with school random effects.

Results shown under Model 1 in Table 3 are mostly qualitatively consistent with those of Model 3 in Table 2. The most important difference, as expected from the gains in efficiency with the joint equation estimation method, is that a few incentive variables that are statistically significant in the joint estimation of Model 3 in Table 2 become statistically insignificant. These include whether teachers can choose their teaching methods in the tutoring equation, whether teachers can develop teaching materials, and whether the school principal can dismiss teachers or evaluate their performance, and the distance to the district education bureau in the work-hours equation. The only exception is salary delay, which is statistically significant in the single equation estimation but insignificant in the joint equation framework.

If we assume that these estimates are asymptotically consistent, an interesting implication related to the effort tradeoff discussed in the previous section is that, where the school principal has the power to dismiss a teacher, teachers are more likely to increase total effort. This is indicated by the strong statistical significance for this variable in the tutoring regression, and the statistical insignificance in the work hours regression (Model 1, Table 3).

[Table 3 about here]

Potential over-reporting of work hours

We argued in previous sections that potential over-reporting of work hours is not of serious concern because there are no set rules about those. Here we use a two-pronged approach to provide a more formal robustness (and heterogeneity) check on this assumption. We estimate our model again, first, with a sample that leaves out five percent of the teachers with the longest work hours, and second, with a sample that excludes those teachers who were reported to be absent three weeks or more (about two percent of all teachers). The first sample may include the very hard workers but can also include those who may inflate their work hours, and the second sample may include those who may be habitual shirkers who might over-report their work hours to compensate for the absences; regardless of the exact mechanism, these two samples can strongly influence our estimation results.

The resulting estimates (shown in Model 2 and Model 3, Table 3) are mostly qualitatively similar to our previous estimates, but two changes are noteworthy in Model 2. The most important change relates to salary delays: its effect becomes negative and marginally statistically significant in the tutoring equation, but positive and highly statistically significant in the work-hours equation. Faced with salary delays, most teachers appear to substitute effort away from tutoring lessons and to more work hours, a choice we are hard pressed to explain.[27] The distance to the district education bureau also loses its statistical significance in the trimmed samples. However, estimation results remain almost unchanged under the complete Model 3, suggesting that these results are not strongly influenced by teachers who were absent more frequently than others.

[27] An explanation is that, faced with longer salary delay, teachers are disencouraged from exerting efforts on the more observable activity of tutoring provision; however, being mostly endowed with high morale (Table 1), they channel their energies instead into the less observable activity of lesson preparation, which is also beneficial for their students.

A different way of aggregating the non-pecuniary incentives

From a policy perspective, it is useful to look at the impact on teacher effort of each non-pecuniary variable. However, since there are several such variables, there is likely to be a multi-collinearity problem if these variables reflect just different aspects of the same enabling environment in a school. It could be useful then to aggregate the two types of non-pecuniary incentive mechanisms. For this purpose, we create a teacher autonomy index and a monitoring index by adding up the corresponding variables with the former ranging from 0 to 3 and the latter from 0 to 4.[28] A higher value for either index indicates a higher level of teacher autonomy or monitoring in the school.

We estimate again Model 3 in Table 2 and show estimation results of Model 4 in Table 3. Again, results are consistent with our two main results in Model 3 in Table 2: higher levels of teacher autonomy and monitoring induce teachers to substitute effort away from the unobservable activity of after-school work hours into the more observable activity of providing tutoring lessons.

Other hetoregenity analyses

Teachers who are civil servants may behave differently from teachers on shorter-term contracts. Studies on India (Banerjee et al., 2007) and Kenya (Duflo, Dupas, and Kremer, 2012) find that short-term contract teachers hired by the local community work harder and more efficiently than their civil-service colleagues. This could be because contract teachers are likely to be subject to more local monitoring and face greater pressure to work harder. Since around one percent of teachers in our sample are hired by the local community, to examine the robustness of results against this hypothesis we exclude from our estimation sample those

[28] Another way to aggregate the variables is to use, say, the first component obtained by the principal component method. We also tried this and found that estimation results are very similar. Thus to make results easier to interpret, we use the simpler arithmetic aggregation instead.

teachers who are posted to their school by the local community (about five percent of all teachers). If the local community can influence the assignment of teachers, then they are also capable to influence decisions such as firing and demotion. Our estimation results using each sample (not shown) provide very qualitatively similar results.

We also implemented other heterogeneity analyses including restricting the estimation sample to schools that exempted less than half of their students from paying tuition fees or adding village characteristics for more control variables.[29] The hypotheses behind these specifications are respectively that teachers at richer schools may exert more effort due to heightened monitoring power from students' parents (see, e.g., World Bank, 2003) or that teacher effort can be influenced by their socio-economic environment and other economic opportunities that can compete for their time away from teaching. Estimation results (not shown) are, however, qualitatively similar.

VI. Further Investigation of After-School Tutoring

Our school survey has a rich design which also offers details about whether teachers provide tutoring for pay or without pay, as well as the number of students that they tutor. We exploit these data for further analysis in this section. Given the rather complicated modeling (and execution) of a simultaneous equation framework if we are to combine this addition information with the working hours equation,[30] we now revert to using single-equation estimation methods. Assuming our estimation sample is large enough, estimates would be consistent but less efficient than the related results in Table 2.

[29] These village characteristics include the share of fathers with lower secondary education or higher in the village, whether the village has a school at the lower secondary or higher level, the availability of non-agricultural work, the average daily wages, and the share of poor households in the village.
[30] For example, a good candidate for a joint estimation of tutoring with or without pay and work hours is the multinomial probit-OLS model with school random effects, which we leave for future research.

VI.1. Tutoring with and without Pay

Teachers who charge a fee for their after-school tutoring may have different incentives from teachers that do not. We thus look into the determinants of teacher provision of tutoring for pay or not. Without assuming that effort is greater or less when tutoring happens for a fee, we use a multinomial logit model with school random effects.[31] The log likelihood function for the entire sample in the multinomial logit model with school random effects is defined as follows

$$\ln L\left(T_{ij} \mid X_{ij}, c_i; \beta, \sigma_c, \rho\right) = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=0}^{K} d_{ijk} \ln P(T_{ij} = k \mid X_{ij}]$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=0}^{K} d_{ijk} \ln\left(\frac{e^{\beta_k X_{ij} + c_i + \varepsilon_{ij}}}{1 + \sum_{k=0}^{K} e^{\beta_k X_{ij} + c_i + \varepsilon_{ij}}} \right), \quad k = 0, 1, 2 \qquad (4)$$

where T_{ijk} indicates the tutoring decision which can be either no tutoring (k= 0), tutoring with pay (k= 1), or tutoring without pay (k= 2), and d_{ijk} is a dummy variable for the choice. The vector of the observables X_{ij} and the school random effects c_i are defined as in equation (2).

Estimation results are provided in Table 4; similar to Table 2, three models with sequential blocks of variables are estimated, and the base category is teachers not offering any after-school tutoring. The variance of the school random effects c_i is highly statistically significant in all models, suggesting that the school random effects component should be included in our estimation models. Our preferred model is Model 3, the most complete model. As expected, fewer variables are statistically significant in this single-equation method than in Table 2 where a joint estimation method is used; but the variables that are statistically significant are similar to the results in Table 2. These include the monitoring-related incentives, whether the school has a

[31] We estimate this model using the user-written command “gllamm” in Stata (Rabe-Hesketh and Skrondal, 2012).

PTA and whether the principal can dismiss teachers. The variables that are statistically significant in Table 4 are only salary delay and whether teachers can adapt the curriculum to the local conditions, with the former being significant only for tutoring without pay and the latter being marginally significant only for tutoring with pay.

[Table 4 about here]

For easier interpretation of the results, the changes in the odds ratios for the estimated coefficients in Model 3 are also provided.[32] The existence of a PTA can increase the odds of teachers providing paid after-school tutoring by a factor of 13.5 and unpaid after-school tutoring by a factor of 76.1, relative to no tutoring. An important change is that salary delay has a stronger and more statistically significant impact on unpaid tutoring: A one-month salary delay reduces the odds of teachers providing free tutoring by 40 percent.

The null hypothesis that the estimated coefficients for the two categories of teachers (i.e. those who tutor for pay and those who tutor without pay) in Model 3 are the same is strongly rejected for (i.e., χ^2 value of 111.85 for 26 degrees of freedom). However, while the level of statistical significance is different for most of the teacher autonomy and monitoring variables for the two categories of teachers, the estimated coefficients on the statistically significant variables are qualitatively similar.

VI.2. Number of Students Tutored

We next turn to the determinants of the number of students that teachers tutor. Since the number of students being tutored is a count variable, we estimate a random effects negative binomial model. This model allows for the large number of zeros for tutored students by its

[32] The odds ratios are obtained by exponentiating the corresponding coefficients. See, for example, Long (1997) for more discussion on the multinomial logit model.

dispersion parameters. [33] The log likelihood function for the entire sample in the random effects negative binomial model is defined as follows

$$\ln L\left(T_{ij} \mid X_{ij}, r, s; \beta, \sigma_c, \rho\right) = \sum_{i=1}^{I} \ln\left(\int_0^{\infty} \prod_{j=1}^{J} P(T_{ij} = k_{ij} \mid X_{ij}, \delta_i) f(\delta_i) d\delta_i\right)$$

$$= \sum_{i=1}^{I} \ln\left(\frac{\Gamma(r+s)\Gamma\left(r + \sum_{j=1}^{J} \lambda_{ij}\right)\Gamma\left(s + \sum_{j=1}^{J} k_{ij}\right)}{\Gamma(r)\Gamma(s)\Gamma\left(r + s + \sum_{j=1}^{J} \lambda_{ij} + \sum_{j=1}^{J} k_{ij}\right)} \prod_{j=1}^{J} \frac{\Gamma(\lambda_{ij} + k_{ij})}{\Gamma(\lambda_{ij})\Gamma(k_{ij}+1)}\right) \quad (5)$$

where k_{ij} indicates the number of students that teacher *i* in school *j* tutors. Assume that $T_{ij}|\gamma_{ij}$ ~Poisson(γ_{ij}), where $\gamma_{ij}|\delta_i$ ~gamma(λ_{ij}, δ_i) with $\lambda_{ij} = e^{\beta x_{ij}}$; and δ_i is the dispersion parameter that can vary across school and $\frac{1}{1+\delta_i} \sim Beta(r,s)$. The vector of the observables X_{ij} is defined in a similar way to those in equations (2).

Estimation results are provided in Table 5, where our preferred model for interpretation is Model 3. Controlling for other factors, both pecuniary and non-pecuniary incentives strongly influence the number of students tutored. The incentives that are statistically significant include salary delay, whether teachers can adapt the curriculum to the local conditions and whether teachers can set standards for student promotion, the existence of PTA, distance to the district education bureau and whether the principal can dismiss teachers For example, keeping other variables constant, a one-month salary delay reduces the number of students whom teachers tutor by a factor of 0.7 (or equivalently by 30 percent), and the existence of a PTA increases the number of students teachers tutor by a factor of 5.1. These results are qualitatively consistent with the previous estimates.

---

[33] See, for example, Hilbe (2011) for more detailed discussion of the negative binomial model.

[Table 5 about here]

VII. Further Investigation of the PTAs

Estimation results have consistently pointed to the positive impacts of a PTA on teacher effort; however, these results look at the difference between schools with a PTA and schools without a PTA. Since 87 percent of all primary schools in Lao PDR have a PTA (Table 1), it would also be useful to understand whether the different characteristics of PTAs can affect teacher effort within this group of schools. Put differently, if parental participation in the education process—just as teacher effort—can be regarded as an unobserved but continuous latent variable that is observed to be zero below a certain threshold (i.e., no PTA) and assumes different values when it is higher than this threshold (i.e., different levels of PTA activities), it is informative to consider the impacts of different measures of PTA activities on teacher effort beyond the dummy variable for the existence of a PTA. For example, are older and more established PTAs more effective than the new ones? Do more active PTAs bring about better results than less active PTAs?

Estimation results for PTA characteristics are provided in Table 6; coefficients on the relevant PTA characteristics are shown in each cell for teacher provision of after-school tutoring and work hours (columns 3 and 4) and the number of tutored students (column 5).[34] The results suggest that the age of the PTA does not make a difference, but parents' level of involvement, as measured by the proportion of parents serving on the PTA, frequency of PTA meetings, and topics of discussion by the PTA do increase teacher effort. A higher proportion of parents serving on the PTA has a positive impact on tutoring as well as on the number of students being

[34] We attempted to estimate the school random effects Probit-OLS in the first two columns with MSL method but found it harder to get these models to converge. This specially applies to the PTA characteristics that have no statistically significant correlation with teacher outcomes. We thus used Gauss-Hermite quadrature instead since these can provide similar qualitative insights into the mechanism behind PTA characteristics and teacher effort.

tutored (row 3). PTAs that meet more often (row 4) decreases tutoring but increases teacher after-school work hours.

Finally, if the most frequently discussed topics in the PTA are related to school facilities, budget and expenses, and fund raising, teachers are more likely to provide tutoring but do not increase number of students tutored or after-school work hours (row 5). When the most frequently discussed topics are instructional methods, teacher promotion and pay, and student performance, teachers are more likely to increase their after-school work hours and the number of tutored students, but the likelihood of teachers offering tutoring does not increase.

[Table 6 about here]

Interestingly enough, independent of our study findings, the Lao Government has been implementing school-based management with features related to finance and governance in a number of provinces (World Bank, 2009). While rigorous impact evaluation of this reform remains to be implemented, our estimation results, as well as studies from other parts of the world, provide supporting evidence for the reform. Enhanced community participation in monitoring and governance at the school level tends to improve service delivery and quality of schools in India (Banerjee et al., 2010) and health clinics in Uganda (Bjorkman and Svensson, 2009), as well as in a number of other developing countries (Bruns, Filmer and Patrinos, 2011).

VIII. Summary

Teacher effort and performance affect student performance, one way or another. In turn, teacher effort is shaped by a variety of factors including the incentives that are embedded in the education system and those that are particular to a school. In this paper, we examine the influence of pecuniary and non-pecuniary incentives that vary across schools using data from a special school survey in Lao PDR that was linked to a national household consumption and expenditures survey. Delays in salary payment, our measure of a pecuniary incentive, affect teacher effort, especially in the form of after-school tutoring. The non-pecuniary incentives that promote teacher effort, particularly teachers' willingness to offer after-school tutoring, include giving teachers more autonomy in choosing their teaching methods, developing teaching materials, and setting standards for student promotion. When there are stronger monitoring mechanisms in the school, teachers tend also to substitute from the unobservable activity of preparing for classes at home towards the more observable activity of providing after-school tutoring. Teachers are more likely to tutor or tutor more students in schools that have a PTA and in schools where the school principal has dismissal authority.

Although our results are consistent across several estimation models, we do not want to overstate them. First, our measures of teacher effort are not derived from direct observations; the data on after-school work hours for preparing classes and after-school tutoring, whether for pay or not, are self-reported. We argue above that while self-reported teacher absences are known to be unreliable, hours of work after school and after-school tutoring are less susceptible to deliberate misreporting because there are no officially mandated hours for these activities. We also find that that a large proportion of teachers report only a small number of after-school work hours rather than the other way around (Figure 1), lending some credence to the data on teacher effort.

Second, we rely on non-experimental data in this paper to estimate the effect of incentives on teacher effort. This approach may be subject to omitted variable bias. To mitigate such bias, we estimate a simultaneous OLS-probit model with school random-effects, which is then supplemented with further analysis of different after-school tutoring as well as PTA activities. Likewise, we control for a set of teacher and school characteristics in different model specifications, as well as apply a number of robustness checks to estimation results. While experimental design studies can provide the benchmark results for the impact of specific interventions, our paper offers a couple of advantages. One advantage is that we use nationally representative survey data for a large sample size from around the country.[35] Another advantage is that while a number of experimental studies are designed to vary certain features of the school system one at a time (for example, providing either more textbooks or additional contract teachers), we are able to examine the overall impact of all relevant components. See, for example, Ravallion (2009) for a thoughtful discussion of these issues. However, a combination of a study using nationally representative data and evaluations of smaller-scale experiments may perhaps offer the best of both worlds.

Finally, we do not have data on student performance such as test scores and thus cannot investigate the impact of incentives and teacher effort on student achievement, the ultimate variable of interest. Here we rely on the insights from previous studies. Hanushek and Rivkin (2010) find that teacher quality (as measured by a teacher fixed-effects component) increases student reading and math achievement by between 0.11 and 0.15 standard deviations in the United States. A study that is closer to ours finds that a 5-percent increase in the teacher's absence rate reduces student learning by 4 to 8 percent of average gains over the year in Zambia (Das et al., 2007).

Our results reveal that teachers engage in strategic choices about effort. Some incentives lead teachers to provide more (observable) after-school tutoring but fewer (unobservable) after-school work hours spent in preparing for classes, and some incentives lead to the opposite choices. To avoid inducing strategic choices that reduce overall effort and lower learning gains, understanding how different forms of pecuniary and non-pecuniary incentives affect teacher

[35] See, for example, Bold et al. (2012) for a discussion on the difficulties of scaling up a randomization study.

effort is critical and should inform the mix of incentive mechanisms used in the school system. Ignoring the differential impacts of, say, greater teacher autonomy and monitoring mechanisms may produce the undesirable (and unanticipated) effect of lowering teacher effort.[36] Moreover, further differentiation of incentives can reveal more. For example, we find that considering the issues that PTAs focus on (such as if they pay more attention to "deeper" issues such as instructional methods show that teachers may prefer to spend more hours on after-school activities that are less visible to monitors but might produce better teaching.

Promising avenues for future research include more investigation into the strategic behavior of teachers under different incentive mechanisms. To do this, better measures of time spent in activities that are not directly observable to PTAs and the principal but have positive impact on student learning, such as hours spent preparing lessons, are needed. Moreover, because of measurement issues related to self-reported effort, direct observation methods akin to those used in time-allocation studies—although these methods are admittedly costly to implement—would help strengthen the quality of future studies. Lastly, following the results chain to the desired end of more learning, the impact of teacher effort on student performance is undoubtedly multi-faceted and may require learning measures that do not rely on single performance test scores (Koretz, 2002).

---

[36] See, for example, Benabou and Tirole (2006) for a literature review and evidence.

References

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis. (2012). "Recent Developments in the Economics of Time Use". *Annual Review of Economics*, 4:373–397.

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

Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. (2010). "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India". *American Economic Journal: Economic Policy*, 2:1, 1–30.

Banerjee, Ritwik, Elizabeth M. King, Peter F. Orazem, and Elizabeth M. Paterno. 2012. "Student and Teacher Attendance: The Role of Shared Goods in Reducing Absenteeism," *Economics of Education Review* 31 (5): 563-574

Benabou, Roland, and Jean Tirole. (2006). "Incentives and Prosocial Behavior." *American Economic Review*, 96(5): 1652–78.

Benveniste, Luis, Jeffery Marshall , Lucrecia Santibañez. (2008). *Teaching in Lao PDR*. Human Development Sector, East Asia and the Pacific Region, World Bank, and Ministry of Education, Lao People's Democratic Republic.

Biswal, Bagala P. (1999). "Private Tutoring and Public Corruption: A Cost-effective Education System for Developing Countries." *Developing Economies*, 37 (2): 222–40.

Bjorkman, Martina, and Jakob Svensson. (2009). "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *Quarterly Journal of Economics*, 124(2): 735–69.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a and Justin Sandefury. (2012). "Interventions and Institutions: Experimental Evidence on Scaling up Education Reforms in Kenya". Preliminary draft. Retrieved from http://www.iies.su.se/polopoly_fs/1.101632.1348137980!/menu/standard/file/2012-08-06%20Kenya%20RCT.pdf (September 2013).

Bray, Mark. (2009). "*Confronting the Shadow Education System. What Government Policies for What Private Tutoring*." Paris: United Nations Education, Scientific and Cultural Organization, International Institute for Educational Planning.

Bray, Mark and Chad Lykins. (2012). *Shadow Education: Private Supplementary Tutoring and Its Implications for Policy Makers in Asia*. Hong Kong: Asian Development Bank and University of Hong Kong.

Bruns, Barbara, Deon Filmer and Harry Anthony Patrinos. (2011). *Making Schools Work New Evidence on Accountability Reforms*. Washington DC: World Bank.

Butler, J. S. and Robert Moffitt. (1982). "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model". *Econometrica*, 50(3): 761-764.

Cappellari, Lorenzo, and Stephen P. Jenkins. (2006). "Calculation of Multivariate Normal Probabilities by Simulation, with Applications to Maximum Simulated Likelihood Estimation". *Stata Journal*, 6(2): 156- 189.

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

Dang, Hai-Anh. (2007). "The Determinants and Impact of Private Tutoring Classes in Vietnam", *Economics of Education Review* 26(6): 684-699.

———. (2008). "*Private Tutoring in Vietnam: An Investigation of its Causes and Impacts with Policy Implications*". VDM Verlag Dr. Mueller Publishing House: Saarbrucken, Germany.

Dang, Hai-Anh, and Halsey Rogers. (2008). "The Growing Phenomenon of Private Tutoring: Does It Deepen Human Capital, Widen Inequalities, or Waste Resources?" *World Bank Research Observer* 23(2): 161-200.

Das, Jishnu, Stefan Dercon, James Habyarimana and Pramila Krishnan. (2007) "Teacher Shocks and Student Learning", *Journal of Human Resources*, 42(4): 820–862.

Di Gropello, Emanuela and Jeffery H. Marshall. (2005). "Teacher Effort and Schooling Outcomes in Rural Honduras". In Emiliana Vegas (eds.), *Incentives to Improve Teaching Lessons: from Latin America*. Washington DC: World Bank.

Dixit, Avinash. (2002). "Incentives and Organizations in the Public Sector: An Interpretative Review". *Journal of Human Resources*, 37(4): 696-727.

Duflo, Esther, Rema Hanna, and Stephen Ryan. (2012). "Incentives Work: Getting Teachers to Come to School". *American Economic Review*, 102(4): 1241–1278.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. (2012). "*School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experiment Evidence from Kenyan Primary Schools*." National Bureau of Economic Research, Working Paper 17939.

Education Statistics and Information Technology Center (ESITC). (2011). *Education Yearbook 2008-2009*. Vientianne, Lao PDR: Ministry of Education.

Glewwe, Paul, and Michael Kremer. (2006). "School, Teachers, and Education Outcomes in Developing Countries." In Eric A. Hanushek and Finis Welch. (Eds.). *Handbook of the Economics of Education*. Amsterdam: North-Holland.

Glewwe, Paul, Nauman Ilias, and Michael Kremer. "Teacher Incentives." *American Economic Journal: Applied Economics*, 2: 205–227

Glewwe, Paul, Alaka Holla and Michael Kremer. (2009). ''Teacher Incentives in the Developing World''. In Mathew G. Springer (Eds.). *Performance Incentives: Their Growing Impact on American K-12 Education.* Washington DC: Brookings Institution Press.

Glick, Peter. (1999). "Simultaneous Determination of Home Work and Market Work of Women in Urban West Africa". *Oxford Bulletin of Economics and Statistics*, 61(1): 57-84.

Gould, William, Jeffrey Pitblado, and William Poi. (2010). "*Maximum Likelihood Estimation with Stata*". Stata Press: Texas.

Gueorguieva, Ralitza V. and Alan Agresti, A. (2001), ''A Correlated Probit Model for Joint Modelling of Clustered Binary and Continuous Responses,'' *Journal of the American Statistical Association*, 96, 1102–1112.

Hanushek, Erik A. and Steven G. Rivkin. (2010). "Generalizations about Using Value-Added Measures of Teacher Quality". *American Economic Review,* 100: 267–271.

Hilbe, Joseph M. (2011). "*Negative Binomial Regression*". United Kingdom: Cambridge University Press.

Ireson, Randall W. (1996). "Invisible Walls: Village Identity and the Maintenance of Cooperation in Laos". *Journal of Southeast Asian Studies,* 27(2): 219-244.

Jacob, Brian A., and Lars Lefgren. (2004). "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis." *Review of Economics and Statistics* 86 (1): 226–44.

---. (2008). "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education". *Journal of Labor Economics*, 26(1): 101-136.

Jayachandran, Seema. (2013). "Incentives to Teach Badly? After-School Tutoring in Developing Countries". Working paper. Department of Economics, Northwestern University.

Kimmel, Jean and Rachel Connelly. (2007). "Mothers' Time Choices Caregiving, Leisure, Home Production, and Paid Work". *Journal of Human Resources,* 42(3): 643-681.

Koretz, Daniel M. (2002). "Limitation in the Use of Achievement Tests as Measures of Educators' Productivity". *Journal of Human Resources*, 37(4): 752-777.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. (2009). "Incentives to Learn". *Review of Economics and Statistics*, 91(3): 437–456.

Kremer, Michael, Conner Brannen, and Rachel Glennerster. (2013). "The Challenge of Education and Learning in the Developing World". *Science*, 340:297-300.

Lavy, Victor. (2009). "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics". *American Economic Review*, 99(5): 1979–2011

Lavy, Victor, and Analia Schlosser. (2005). "Targeted Remedial Education for Underperforming Teenagers: Costs and Benefits." *Journal of Labor Economics,* 23(4): 839–74.

Long, J. Scott. (1997). "*Regression Models for Categorical and Limited Dependent Variables"*. Thousand Oaks, California: SAGE Publications, Inc.

Liu, Xuefeng, Michael J. Daniels, and Bess Marcus. (2009). "Joint Models for the Association of Longitudinal Binary and Continuous Processes With application to a Smoking Cessation Trial". *Journal of the American Statistical Association*, Vol. 104: 429- 438.

Neal, Derek. (2011). "The Design of Performance Pay in Education". In Eric A. Hanushek Stephen Machin and Ludger Woessmann. (Eds.). *Handbook of the Economics of Education*. Amsterdam: North-Holland.

Podgursky, Michael. (2011). "Teacher Compensation and Collective Bargaining". In Eric A. Hanushek Stephen Machin and Ludger Woessmann. (Eds.). Handbook of the Economics *of Education*. Amsterdam: North-Holland.

Podgursky, Michael J. and Matthew G. Springer. (2007). "Teacher Performance Pay: A Review". *Journal of Policy Analysis and Management*, Vol. 26(4): 909–949.

Rabe-Hesketh, Sophia, and Anders Skrondal. (2012). "*Multilevel and Longitudinal Modeling Using Stata. Volume 1: Continuous Responses*". College Station, TX: Stata Press.

Ravallion, Martin. (2009). "Should the Randomistas Rule?" *Economist's Voice*, 6(2): 1-5.

Rigg, Jonathan. (2009). "A Particular Place? Laos and its Incorporation into the Development Mainstream". *Environment and Planning A*, 41: 703-721.

Sándor, Zsolt and Péter András. (2004). "Alternative Sampling Methods for Estimating Multivariate Normal Probabilities". *Journal of Econometrics*, 120(2): 207-234.

Sawada, Yasuyuki and Andrew B. Ragatz.(2005). "Decentralization of Education, Teacher Behavior, and Outcomes: The Case of El Salvador's EDUCO Program". In Emiliana Vegas. (Eds.). *Incentives to Improve Teaching Lessons: Lessons from Latin America*. Washington DC: World Bank.

Silova, Iveta, and Mark Bray. (2006). "The Hidden Marketplace: Private Tutoring in Former Socialist Countries." In Iveta Silova, Virnigija Budiene, and Mark Bray, (Eds.). *Education in a Hidden Marketplace: Monitoring of Private Tutoring*. New York: Open Society Institute.

Taylor, Eric s. and John Tyler. (2012). "The Effect of Evaluation on Teacher Performance". *American Economic Review*, 102(7): 3628–3651.

Train, Kenneth. (2009). "*Discrete Choice Methods with Simulation*". Cambridge University Press: United Kingdom.

Vaenkeo, Souksakhone. (2010). "Teachers Frustrated with Delays in Salary Payments". *Vientiane Times*, April 9, 2010.

Waterreus, Ib and Simone Dobbelsteen. (2001). "Wages and Teachers' Hours of Work". *De Economist*, 149(3): 277–298.
Wooldridge, Jeffrey M. (2010). "*Econometric Analysis of Cross Section and Panel Data*". Cambridge, Mass.: MIT Press.
World Bank. (2003). *World Development Report 2004: Making Services Work for Poor People*. Washington DC: World Bank.
---. (2008). "*Lao PDR Public Expenditure Tracking Survey in Primary Education and Primary Health*". Poverty Reduction and Economic Management Sector Unit East Asia and Pacific Region.
---. (2009). *Lao People's Democratic Republic - Education for All-Fast Track Initiative (EFA-FTI) Catalytic Fund Application Project*. Washington D.C.: Worldbank. Retrieved from http://documents.worldbank.org/curated/en/2009/08/11168474/lao-peoples-democratic-republic-education-all-fast-track-initiative-efa-fti-catalytic-fund-application-project (September 2013).
---. (2013). *World Development Indicators Online*. Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators (September 2013).
Yamada, Hiroyuki, Yasuyuki Sawada, and Xubei Luo. (2013). "Why is Absenteeism Low among Public Health Workers in Lao PDR?" *Journal of Development Studies*, 49(1): 125-133.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Obs.
Teacher Effort					
Teacher offers tutoring	0.13	0.33	0	1	1532
Teacher offers tutoring with pay	0.05	0.21	0	1	1532
Teacher offers tutoring without pay	0.08	0.27	0	1	1532
Number of students tutored	1.10	5.23	0	63	1447
Teachers' after-school work hours	12.56	10.04	0	50	1532
Teacher Incentive/ Autonomy					
No of months salary in arrears	1.77	1.70	0	7	1532
Teacher can choose teaching method	0.31	0.46	0	1	1532
Teacher can develop teaching materials	0.41	0.49	0	1	1532
Teacher can adapt curriculum to local conditions	0.14	0.35	0	1	1532
Teacher can set standards for student promotions	0.27	0.44	0	1	1532
Monitoring & Principal Management Power					
School has PTA	0.87	0.34	0	1	1532
Log of distance to district edu. bureau	2.06	1.30	0	5.30	1532
Principal can dismiss teachers	0.09	0.29	0	1	1532
Principal can evaluate teacher performance	0.79	0.41	0	1	1532
Principal can set teachers' working hours	0.24	0.43	0	1	1532
Teacher characteristics					
Female	0.52	0.50	0	1	1532
Lao-Tai	0.86	0.34	0	1	1532
Born in this district	0.72	0.45	0	1	1532
Upper secondary education	0.06	0.25	0	1	1532
Vocational education	0.79	0.41	0	1	1532
College education	0.03	0.17	0	1	1532
Teaching experience	13.88	9.04	0	63	1532
Permanent position	0.93	0.25	0	1	1532
Teacher has a guidebook	0.97	0.17	0	1	1532
Number of grades	1.32	1.00	0	9	1532
Number of students	30.43	13.77	0	69	1532
School characteristics					
School has multigrade teaching	0.19	0.39	0	1	1532
School fee index	3.42	2.42	0	12	1532
Log of distance to provincial capital	3.23	1.60	0	5.91	1532
Log of distance to nearest paved road	1.54	1.78	0	6.40	1532
Urban area	0.34	0.47	0	1	1532

Table 2: Determinants of After-School Tutoring and Work Hours, Joint OLS-probit Model

	Model 1		Model 2		Model 3		
	Tutoring	Working Hours	Tutoring	Working Hours	Tutoring	Working Hours	Mar. effects
Teacher Incentive/ Autonomy							
No of months salary in arrears	-0.016	-0.265	-0.013	-0.225	-0.035	-0.152	
	(-0.14)	(-1.55)	(-0.13)	(-1.37)	(-0.32)	(-0.92)	
Teacher can choose teaching method	2.129***	-4.260***	2.264***	-4.171***	2.260***	-4.069***	0.11 (d)
	(2.63)	(-4.94)	(2.86)	(-5.01)	(2.80)	(-4.82)	
Teacher can develop teaching materials	-0.638	2.858***	-0.364	3.064***	-0.695	2.886***	
	(-0.84)	(3.49)	(-0.40)	(3.54)	(-0.91)	(3.47)	
Teacher can adapt curriculum to local conditions	-0.550	-1.072	-0.376	-1.006	-0.103	-0.989	
	(-0.67)	(-1.10)	(-0.51)	(-1.08)	(-0.11)	(-0.95)	
Teacher can set standards for student promotions	1.033	-0.578	1.192	-0.619	0.944	-0.761	
	(1.34)	(-0.72)	(1.48)	(-0.77)	(1.59)	(-1.01)	
Monitoring & Principal Management Power							
School has PTA	2.423***	1.466	2.505***	1.416*	2.843***	0.870	0.12 (d)
	(2.61)	(1.64)	(3.03)	(1.66)	(3.04)	(0.96)	
Log of distance to district edu. bureau	-0.124	0.356	-0.241	0.234	-0.285	0.733**	
	(-0.57)	(1.22)	(-1.15)	(0.82)	(-1.07)	(2.25)	
Principal can dismiss teachers	3.858***	-1.057	3.850***	-1.458	4.911***	-2.186*	0.34 (d)
	(3.37)	(-0.83)	(4.29)	(-1.30)	(4.86)	(-1.86)	
Principal can evaluate teacher performance	-0.062	-1.731**	-0.116	-2.046***	0.729	-2.119**	
	(-0.09)	(-2.12)	(-0.19)	(-2.61)	(0.81)	(-2.52)	
Principal can set teachers' working hours	0.141	-0.456	-0.646	-0.887	-0.567	-0.512	
	(0.18)	(-0.55)	(-0.90)	(-1.13)	(-0.42)	(-0.50)	
Teacher characteristics							
Female	-0.402**	1.073**	-0.460***	0.833*	-0.439**	0.658	-0.02 (d)
	(-2.18)	(2.14)	(-2.60)	(1.71)	(-2.52)	(1.37)	
Lao-Tai	-0.917*	-0.139	-0.899**	0.051	-0.663	-0.121	
	(-1.95)	(-0.17)	(-2.02)	(0.06)	(-1.44)	(-0.15)	
Born in this district	-0.012	-0.054	-0.040	-0.108	-0.093	0.283	
	(-0.07)	(-0.10)	(-0.25)	(-0.21)	(-0.57)	(0.56)	
Upper secondary education	0.869*	1.054	0.995*	1.464	0.961*	1.188	0.04 (d)
	(1.65)	(0.93)	(1.95)	(1.33)	(1.84)	(1.10)	
Vocational education	0.598	1.851**	0.700**	1.707**	0.727**	1.409*	0.03 (d)
	(1.61)	(2.40)	(2.01)	(2.28)	(2.06)	(1.92)	
College education	0.750	-1.801	0.957*	-1.119	0.979*	-1.705	0.05 (d)
	(1.36)	(-1.16)	(1.88)	(-0.75)	(1.75)	(-1.16)	
Teaching experience	0.004	-0.024	0.010	-0.012	0.014	-0.032	
	(0.46)	(-0.90)	(1.05)	(-0.48)	(1.50)	(-1.26)	
Permanent position			-0.108	1.814*	-0.106	1.728*	
			(-0.27)	(1.92)	(-0.27)	(1.88)	
Teacher has a guidebook			0.664*	4.283***	0.624	4.169***	
			(1.65)	(3.23)	(1.48)	(3.22)	
Number of grades			0.118	-0.268	0.096	-0.128	
			(1.39)	(-1.12)	(1.12)	(-0.53)	
Number of students			0.011*	0.123***	0.014**	0.111***	0.001
			(1.74)	(7.17)	(2.09)	(6.62)	
School characteristics							
School has multigrade teaching					-0.109	0.065	
					(-0.09)	(0.06)	
School fee index					-0.014	0.433***	
					(-0.07)	(2.60)	
Log of distance to provincial capital					-0.652***	-0.524**	-0.03
					(-3.09)	(-1.99)	
Log of distance to nearest paved road					0.927***	-0.442**	0.04
					(5.38)	(-2.14)	
Urban area	1.109*	1.899**	1.090**	1.810**	0.592	0.553	
	(1.79)	(2.32)	(1.99)	(2.31)	(0.47)	(0.49)	
Constant	-7.608***	10.953***	-8.746***	2.123	-9.078***	3.655	
	(-5.69)	(6.98)	(-6.50)	(1.01)	(-4.32)	(1.48)	
σ_v		8.33		8.00		7.79	
σ_c	4.50***		4.55***		4.72***		
ρ_school	0.95	0.23	0.95	0.24	0.96	0.27	
ρ_teacher	-0.62***		-0.67***		-0.67***		
χ2	37.88		83.23		108.63		
Log likelihood	-5865.15		-5815.38		-5786.93		
N	1532		1532		1532		

Note: 1. * p<0.1, ** p<0.05, *** p<0.01, z-statistics in parentheses.
2. All models are estimated jointly using the school random-effects Probit-OLS model.
3. The marginal effects in Model 3 are calculated for the tutoring regression; (d) stands for discrete change of dummy variables from 0 to 1.
4. ρ_school is the unconditional intra-school correlation coeffient; ρ_teacher is the conditional intra-class correlation coeffient for the same teacher.

Table 3: Robustness Checks on the Determinants of After-School Tutoring and Work Hours, Joint OLS-probit Model

	Model 1		Model 2		Model 3		Model 4	
	Tutoring	Working Hours	Tutoring	Working Hours	Tutoring	Working Hours	Tutoring	Working Hours
Teacher Incentive/ Autonomy								
No of months salary in arrears	-0.212**	0.025	-0.185*	0.392***	-0.025	-0.118	-0.064	-0.051
	(-2.11)	(0.13)	(-1.75)	(2.96)	(-0.23)	(-0.71)	(-0.62)	(-0.31)
Teacher can choose teaching method	0.686	-3.888***	1.065*	-1.938***	2.280***	-4.024***		
	(1.40)	(-2.99)	(1.70)	(-2.95)	(3.03)	(-4.80)		
Teacher can develop teaching materials	0.523	1.871	0.270	2.102***	-0.800	2.703***		
	(1.08)	(1.50)	(0.48)	(3.38)	(-1.33)	(3.54)		
Teacher can adapt curriculum to local conditions	-0.972	0.093	-0.003	0.091	0.054	-0.864		
	(-1.62)	(0.06)	(-0.00)	(0.12)	(0.07)	(-0.91)		
Teacher can set standards for student promotions	-0.226	0.194	-0.151	-0.598	0.776	-0.874		
	(-0.46)	(0.16)	(-0.20)	(-0.84)	(1.38)	(-1.18)		
Monitoring & Principal Management Power								
School has PTA	1.466**	1.281	3.198***	0.195	2.820***	1.166	2.535***	1.407
	(2.17)	(1.08)	(3.81)	(0.27)	(3.25)	(1.30)	(3.18)	(1.62)
Log of distance to district edu. bureau	0.124	0.217	0.084	-0.002	-0.344	0.657**	-0.489*	0.449
	(0.58)	(0.39)	(0.31)	(-0.01)	(-1.50)	(2.06)	(-1.86)	(1.41)
Principal can dismiss teachers	2.274***	-0.494	3.915***	-1.972*	4.982***	-2.037*		
	(2.60)	(-0.24)	(3.38)	(-1.72)	(4.75)	(-1.70)		
Principal can evaluate teacher performance	-0.076	-1.364	0.417	-1.186*	0.803	-1.999**		
	(-0.16)	(-1.17)	(0.67)	(-1.81)	(1.10)	(-2.44)		
Principal can set teachers' working hours	-0.517	0.354	-0.175	0.853	-0.804	-0.596		
	(-1.11)	(0.30)	(-0.29)	(1.35)				
Incentives & Monitoring Indexes							0.707***	-0.469*
Teacher autonomy index							(3.36)	(-1.85)
							0.929**	-1.520***
Principal management power index							(2.05)	(-3.11)
Other control variables								
Teacher and school characteristics	Y		Y		Y		Y	
σ_v		4.54	5.70		7.85		8.13	
σ_c	2.79***	8.22***	3.84***		4.70***		4.43***	
ρ_school	0.89	0.77	0.94	0.31	0.96	0.26	0.95	0.23
ρ_teacher			-0.62***		-0.66***		-0.68***	
χ2	37.13	225.09	81.15		97.31		71.14	
Log likelihood	-298.12		-5034.88		-5668.71		-5835.75	
N	1532	1532	1446		1497		1532	

Note: 1. * p<0.1, ** p<0.05, *** p<0.01, z-statistics in parentheses.
2. All models are estimated jointly using the school random-effects Probit-OLS model, except for Model 1 where the Tutoring and Working Hours equations are estimated separately with a school random-effects probit model and a linear school random-effects model respectively. Model 2 excludes five percent of teachers who report longest after-school hours. Model 3 excludes teachers who were absent for three weeks or more in the past school year. Model 4 controls for aggregated monitoring and incentive indexes.
3. Teacher and school characteristics include the same variables as in Table 2. Teacher characteristics include dummy variables for being female, Lao-Tai, being born in this district, completing upper secondary education, completing vocational education, completing college education, whether the teacher has a permanent position and whether the teacher has a guidebook, the number of years working as a teacher, the numbers of grades and students the teacher teaches. School characteristics include dummy variables indicating whether the school offers multigrade teacher and whether the school is in urban area, a school fee index, and logs of the distances to the provincial capital and the nearest paved road.
4. ρ_school is the unconditional intra-school correlation coeffient; ρ_teacher is the conditional intra-class correlation coeffient for the same teacher.

Table 4: Determinants of After-School Tutoring, Multinomial Logit

	Model 1		Model 2		Model 3			
	Tutoring w. pay	Tutoring w/o pay	Tutoring w. pay	Tutoring w/o pay	Tutoring w. pay	Tutoring w/o pay	Tutoring w. pay, change in odds	Tutoring w/o pay, change in odds
Teacher Incentive/ Autonomy								
No of months salary in arrears	-0.338	-0.514**	-0.312	-0.516**	-0.247	-0.513**		0.6
	(-1.39)	(-2.13)	(-1.29)	(-2.18)	(-1.01)	(-2.18)		
Teacher can choose teaching method	1.420	1.292	1.846	1.648	1.534	1.439		
	(0.92)	(0.86)	(1.20)	(1.10)	(0.99)	(0.95)		
Teacher can develop teaching materials	0.145	1.230	-0.046	1.033	0.487	1.294		
	(0.10)	(0.83)	(-0.03)	(0.76)	(0.33)	(0.90)		
Teacher can adapt curriculum to local conditions	-3.253*	-1.344	-3.470**	-1.457	-3.352*	-1.488	0.04	
	(-1.82)	(-0.81)	(-2.00)	(-0.92)	(-1.85)	(-0.91)		
Teacher can set standards for student promotions	1.036	-1.371	0.988	-1.298	0.954	-1.394		
	(0.74)	(-0.97)	(0.74)	(-0.98)	(0.68)	(-0.99)		
Monitoring & Principal Management Power								
School has PTA	2.658*	2.762**	2.912*	2.735**	2.599*	2.941**	13.5	18.9
	(1.86)	(2.05)	(1.94)	(1.99)	(1.75)	(2.12)		
Log of distance to district edu. bureau	-0.468	0.611	-0.608	0.567	-0.390	0.653		
	(-0.83)	(1.11)	(-1.14)	(1.11)	(-0.63)	(1.08)		
Principal can dismiss teachers	5.362**	4.421*	4.788***	3.949**	4.332*	4.168*	76.1	64.6
	(2.33)	(1.94)	(2.84)	(2.39)	(1.79)	(1.74)		
Principal can evaluate teacher performance	-1.020	0.080	-1.110	0.032	-1.440	0.641		
	(-0.74)	(0.06)	(-0.81)	(0.02)	(-1.03)	(0.47)		
Principal can set teachers' working hours	-0.793	-1.758	-0.741	-1.694	-0.714	-1.337		
	(-0.54)	(-1.22)	(-0.51)	(-1.19)	(-0.50)	(-0.97)		
Teacher characteristics	Y		Y		Y			
School-specific teacher characteristics	N		Y		Y			
School characteristics	N		N		Y			
σ_c	5.470***		5.416***		5.229***			
ρ_school	0.90		0.90		0.89			
χ2	72.44		79.40		378.57			
Log likelihood	-393.50		-389.19		-381.64			
N	1532		1532		1532			

Note: 1. * p<0.1, ** p<0.05, *** p<0.01, t-statistics in parentheses.
2. All models are estimated jointly using the school random-effects multinomial logit model.
3. The base category is teachers' offering no private tutoring classes.
4. Teacher, teacher-specific and school characteristics include the same variables as in Table 2. Teacher characteristics include dummy variables for being female, Lao-Tai, being born in this district, completing upper secondary education, completing vocational education, completing college education, and the number of years working as a teacher. School-specific teacher characteristics include dummy variables for whether the teacher has a permanent position and whether the teacher has a guidebook, the numbers of grades and students the teacher teaches. School characteristics include dummy variables indicating whether the school offers multigrade teacher and whether the school is in urban area, a school fee index, and logs of the distances to the provincial capital and the nearest paved road.
5. The χ2 test is for testing the null that the coeffiecients in the two categories are not different from each other.

Table 5: Determinants of Number of Tutored Students, Negative Binomial Model

	Model 1	Model 2	Model 3	
	No of students tutoted	No of students tutoted	No of students tutoted	Factor change
Teacher Incentive/ Autonomy				
No of months salary in arrears	-0.389***	-0.389***	-0.306***	0.7
	(-3.62)	(-3.62)	(-2.65)	
Teacher can choose teaching method	0.089	0.089	0.014	
	(0.20)	(0.20)	(0.03)	
Teacher can develop teaching materials	-0.088	-0.088	-0.042	
	(-0.19)	(-0.19)	(-0.08)	
Teacher can adapt curriculum to local conditions	-1.516**	-1.516**	-1.312*	0.3
	(-2.27)	(-2.27)	(-1.96)	
Teacher can set standards for student promotions	1.321***	1.321***	1.250***	3.5
	(3.01)	(3.01)	(2.66)	
Monitoring & Principal Management Power				
School has PTA	1.323**	1.323**	1.628***	5.1
	(2.44)	(2.44)	(2.71)	
Log of distance to district edu. bureau	-0.429***	-0.429***	-0.407**	0.7
	(-2.61)	(-2.61)	(-2.14)	
Principal can dismiss teachers	2.149***	2.149***	1.948***	7.0
	(3.69)	(3.69)	(3.18)	
Principal can evaluate teacher performance	-0.255	-0.255	-0.443	
	(-0.54)	(-0.54)	(-0.80)	
Principal can set teachers' working hours	-0.179	-0.179	-0.131	
	(-0.34)	(-0.34)	(-0.24)	
Teacher characteristics	Y	Y	Y	
School-specific teacher characteristics	N	Y	Y	
School characteristics	N	N	Y	
r	0.38	0.38	0.37	
s	0.12	0.12	0.12	
χ2	100.95	100.95	79.36	
Log likelihood	-682.75	-682.75	-675.34	
N	1447	1447	1447	

Note: 1. * p<0.1, ** p<0.05, *** p<0.01, t-statistics in parentheses.
2. All models are estimated using the school random-effects negative binomial model.
3. Teacher, teacher-specific and school characteristics include the same variables as in Table 2. Teacher characteristics include dummy variables for being female, Lao-Tai, being born in this district, completing upper secondary education, completing vocational education, completing college education, and the number of years working as a teacher. School-specific teacher characteristics include dummy variables for whether the teacher has a permanent position and whether the teacher has a guidebook, the numbers of grades and students the teacher teaches. School characteristics include dummy variables indicating whether the school offers multigrade teacher and whether the school is in urban area, a school fee index, and logs of the distances to the provincial capital and the nearest paved road.
4. The χ2 test is for testing the null that the dispersion parameter in the random-effects negative binomial model is constant.

Table 6: Effect of PTAs on Teacher Effort

No		Probit-OLS Model		Negative Binomial Model
		Tutoring	Working Hours	No of students tutored
1	Tenure length for PTA members	-0.485***	0.044	-0.310
		(-5.52)	(0.17)	(0.218)
2	Number of years PTA in existence	0.018	0.011	-0.020
		(1.09)	(0.27)	(0.034)
3	Share of parents on PTA	0.011***	-0.007	0.039***
		(2.91)	(-0.83)	(0.009)
4	Meeting frequency for PTA in the school year	-0.294***	0.909***	-0.347**
		(-4.50)	(5.04)	(0.145)
5	PTA frequently discuss school issues	1.546*	0.044	18.245
		(1.78)	(0.03)	(0.01)
6	PTA frequently discuss instructional methods	-1.205***	2.219***	-0.436
		(-5.44)	(3.48)	(0.454)
7	PTA frequently discuss teacher promotions and pay	-0.187	0.241	1.860***
		(-0.84)	(0.39)	(0.472)
8	PTA frequently discuss student performance	-1.171***	0.880	-0.302
		(-4.77)	(1.21)	(0.584)

Note: 1. * p<0.1, ** p<0.05, *** p<0.01, z-statistics in parentheses. Each cell in the first two columns represents estimated coefficents from the school random-effects Probit-OLS model; each cell in the third column represents those from the school random-effects negative binomial model.

2. Meeting frequency for PTA (row 5) is assigned the following values corresponding to different frequency levels of PTA meeting: 1 (occasionally), 2 (yearly), 3 (twice a year), 4 (quarterly), 5 (monthly), 6 (twice a month), and 7 (weekly or more).

3. School issues (row 5) include school budget and expenses, facilities, and fundraising.

Figure 1: Distribution of Number of Teachers' After-School Work Hours

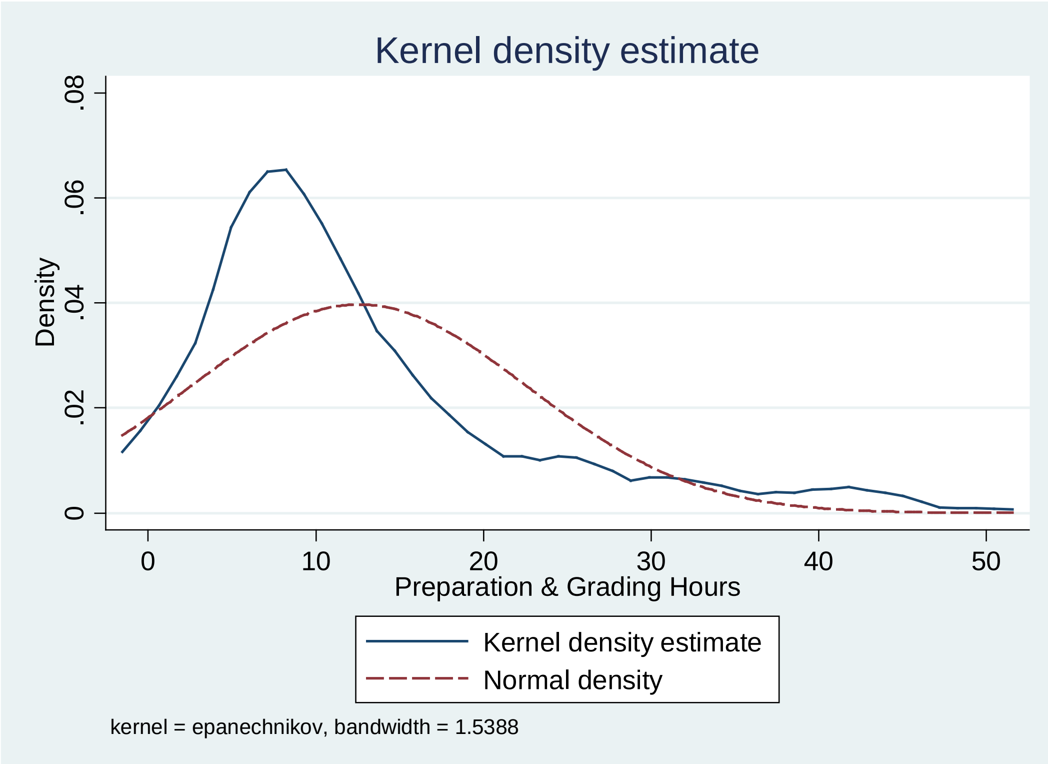
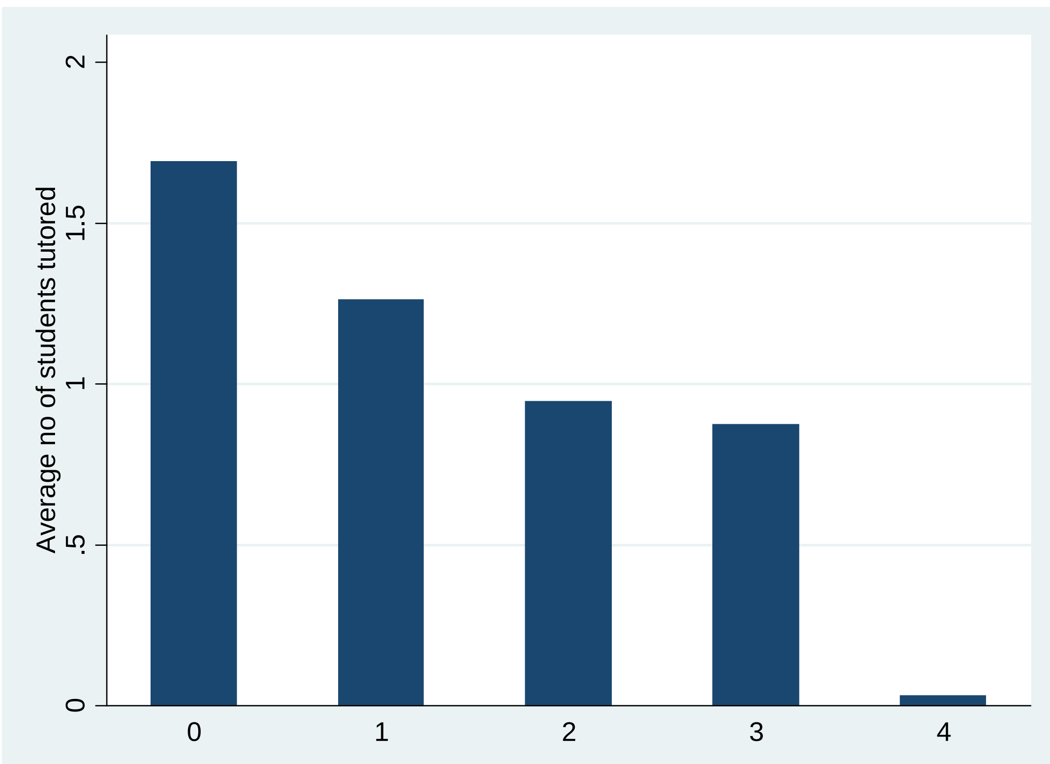


Figure 2: Number of Students Tutored vs. Number of Months of Salary in Arrears



Appendix 1

To derive the likelihood function in a format amenable to estimation, we can rewrite the conditional joint density of (T_{ij}, H_{ij}| X_{ij}, c_i) as the product of two conditional marginal densities

$$f(T_{ij},H_{ij} \mid X_{ij},c_i) = f(T_{ij} \mid H_{ij},X_{ij},c_i)\, f(H_{ij} \mid X_{ij},c_i) \quad \text{(A1)}$$

Under our model assumptions, we have

$$\varepsilon_{ij} = \frac{\rho}{\sigma_\upsilon}\upsilon_{ij} + e_{ij} \quad \text{(A2)}$$

where e_{ij} is a random error term independent of υ_{ij} and Z_{ij}. Given the joint normality of ($\varepsilon_{ij}, \upsilon_{ij}$), it can easily be shown from (A1) that $e_{ij} \mid (X_{ij},\upsilon_{ij},c_i) \sim N(0, 1-\rho^2)$ (*).

Thus combining (A2) and the result in (*), we have

$$P(T_{ij}=1 \mid X_{ij},c_i) = P\left[\beta X_{ij} + c_i + \varepsilon_{ij} \geq 0\right] = \Phi\left[\frac{\beta X_{ij} + \frac{\rho}{\sigma_\upsilon}(H_{ij} - \alpha X_{ij}) + (1 - \frac{\rho}{\sigma_\upsilon})c_i}{\sqrt{1-\rho^2}}\right] \quad \text{(A3)}$$

where $\Phi(.)$ represents the standard cumulative density function.

Since $H_{ij} \mid (X_{ij},c_i) \sim N(0,\sigma_V)$, the second term on the RHS of equation (A1) is simply

$$f(H_{ij} \mid X_{ij},c_i) = \frac{1}{\sigma_\nu}\phi\left[\frac{H_{ij} - \alpha X_{ij} - c_i}{\sigma_\nu}\right] \quad \text{(A4)}$$

where $\phi(.)$ represents the standard density function.

To make notations less cluttered, let $A_{ij} = \dfrac{\beta X_{ij} + \frac{\rho}{\sigma_\upsilon}(H_{ij} - \alpha X_{ij}) + (1 - \frac{\rho}{\sigma_\upsilon})c_i}{\sqrt{1-\rho^2}}$, $B_{ij} = \dfrac{H_{ij} - \alpha X_{ij} - c_i}{\sigma_\nu}$ and $q_{ij} = 2y_{ij} - 1$. Then, plugging (A3) and (A4) into (A1), the likelihood function in (A2) can simply be written as

$$f(T_{ij},H_{ij} \mid X_{ij},c_i) = \Phi(q_{ij}A_{ij})\frac{1}{\sigma_\nu}\phi(B_{ij}) \quad \text{(A5)}$$

Since c_i is assumed to have a normal distribution with mean zero and variance σ_c^2, we can integrate out c_i from (A5). After taking log of the conditional likelihood function for the entire sample we have

$$L\left(T_{ij},H_{ij} \mid X_{ij},c_i;\alpha,\beta,\gamma,\sigma_\upsilon,\sigma_c,\rho\right) = \sum_{i=1}^{I}\log\left(\int_{-\infty}^{\infty}\left[\prod_{j=1}^{J}\Phi(q_{ij}A_{ij})\frac{1}{\sigma_\nu}\phi(B_{ij})\right]\frac{1}{\sigma_c}\phi\left(\frac{c}{\sigma_c}\right)dc\right) \quad \text{(A6)}$$

It is straightforward to see that the integral in (A6) is amenable to simulation. To see this, let $\frac{c}{\sigma_c} = k$, then $c = \sigma_c k$ and $dc = \sigma_c dk$, then we can rewrite the log likelihood function as

$$L\left(T_{ij},H_{ij} \mid X_{ij},c_i;\alpha,\beta,\sigma_\upsilon,\sigma_c,\rho\right) = \sum_{i=1}^{I}\log\left(\frac{1}{S}\sum_{s=1}^{S}\left[\prod_{j=1}^{J}\left[\Phi(q_{ij}A_{ijs})\frac{1}{\sigma_\nu}\phi(B_{ijs})\right]\right]\right) \quad \text{(A7)}$$

with $A_{ijs} = \left[\dfrac{\beta X_{ij} + \dfrac{\rho}{\sigma_\upsilon}(H_{ij} - \alpha X_{ij}) + (1 - \dfrac{\rho}{\sigma_\upsilon})\sigma_c k_s}{\sqrt{1-\rho^2}} \right]$ and $B_{ijs} = \dfrac{H_{ij} - \alpha X_{ij} - \sqrt{2}\sigma_c k_s}{\sigma_\nu}$, where k_s is the sth simulation from a N(0,1) distribution.

To facilitate the maximization of (A7), we usually transform σ_c and σ_υ to the log scale, and ρ as $\rho^* = \dfrac{1}{2}\ln\left(\dfrac{1+\rho}{1-\rho}\right)$. Thus let $P_{is} = \prod_{j=1}^{J}\left[\Phi(q_{ij}A_{ijs})\dfrac{1}{\sigma_\nu}\phi(B_{ijs})\right]$, taking the first derivative of lnL$_i$ with regards to $X_{ij}\beta, Z_{ij}\alpha, \ln\sigma_\upsilon, \ln\sigma_c$ and ρ^* we have

$$\frac{\partial \ln L_i}{\partial \beta X_{ij}} = \frac{\sum_{s=1}^{S}\left\{\dfrac{1}{\sqrt{1-\rho^2}}\dfrac{q_{ij}\phi(q_{ij}A_{ijs})}{\Phi(q_{ij}A_{ijs})}\right\}P_{is}}{\sum_{s=1}^{S}P_{is}} \tag{A8}$$

$$\frac{\partial \ln L_i}{\partial \alpha X_{ij}} = \frac{\sum_{s=1}^{S}\left\{\left[\dfrac{-\rho}{\sigma_\nu\sqrt{1-\rho^2}}\dfrac{q_{ij}\phi(q_{ij}A_{ijs})}{\Phi(q_{ij}A_{ijs})}\right] + \dfrac{B_{ijs}}{\sigma_\nu}\right\}P_{is}}{\sum_{s=1}^{S}P_{is}} \tag{A9}$$

$$\frac{\partial \ln L_i}{\partial \ln \sigma_\upsilon} = \frac{\sum_{s=1}^{S}\left\{\sum_{j=1}^{J}\left[\dfrac{-\rho B_{ijs}}{\sqrt{1-\rho^2}}\dfrac{q_{ij}\phi(q_{ij}A_{ijs})}{\Phi(q_{ij}A_{ijs})}\right] - J + \sum_{j=1}^{J}B_{ijs}{}^2\right\}P_{is}}{\sum_{s=1}^{S}P_{is}} \tag{A10}$$

$$\frac{\partial \ln L_i}{\partial \ln \sigma_c} = \frac{\sum_{s=1}^{S}\left\{\sum_{j=1}^{J}\left[\dfrac{\left(1-\dfrac{\rho}{\sigma_\nu}\right)\sigma_c k_s}{\sqrt{1-\rho^2}}\dfrac{q_{ij}\phi(q_{ij}A_{ijs})}{\Phi(q_{ij}A_{ijs})}\right] + \dfrac{\sigma_c k_s}{\sigma_\nu}\sum_{j=1}^{J}B_{ijs}\right\}P_{is}}{\sum_{s=1}^{S}P_{is}} \tag{A11}$$

$$\frac{\partial \ln L_i}{\partial \rho^*} = \frac{\sum_{s=1}^{S}\left\{\sum_{j=1}^{J}\left[(\sqrt{1-\rho^2}B_{ijs} + \rho A_{ijs})\dfrac{q_{ij}\phi(q_{ij}A_{ijs})}{\Phi(q_{ij}A_{ijs})}\right]\right\}P_{is}}{\sum_{s=1}^{S}P_{is}} \tag{A12}$$