HKUST IEMS
INSTITUTE FOR EMERGING MARKET STUDIES

Promotion Incentives in the Public Sector: Evidence from Chinese Schools

Naureen KARACHIWALLA, Albert PARK

HKUST IEMS Working Paper No. 2015-09

February 2015

HKUST Institute for Emering Market Studies
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
T: +852 3469 2215 E: iems@ust.hk W: iems.ust.hk

Promotion Incentives in the Public Sector: Evidence from Chinese Schools

Naureen KARACHIWALLA, Albert PARK

HKUST IEMS Working Paper No. 2015-09
February 2015

Abstract

We provide the first evidence that promotion incentives can influence effort of employees in the public sector by studying China's system of annual evaluations and promotions for teachers. Theoretical predictions from a tournament model of promotion incentives are tested using panel data on primary and middle school teachers in western China. Consistent with theory, we find that promotions are associated with significant wage increases, that higher wage increases are associated with higher effort, that teachers increase effort in the years leading up to promotion eligibility but reduce effort if they are repeatedly passed over for promotion, and that increasing the number of competitors reduces the relative performance of those at the extremes of the skill distribution. Evaluation scores are positively associated with time spent on teaching and with student test scores, diminishing concerns that evaluations are manipulated.

Author's contact information

Naureen Karachiwalla
International Food Policy Research Institute (IFPRI)
E: N.Karachiwalla@cgiar.org

Albert Park
HKUST Institute for Emerging Market Studies, Division of Social Science, and Department of Economics
The Hong Kong University of Science and Technology
T : +852 2358 5981
E: albertpark@ust.hk

Promotion incentives in the public sector: evidence from Chinese schools*

Naureen Karachiwalla[†] Albert Park[‡]

February 1, 2015

Abstract

We provide the first evidence that promotion incentives influence effort of public employees by studying China's system of promotions for teachers. Theoretical predictions from a tournament model of promotion incentives are tested using panel data on primary and middle school teachers. Consistent with theory, promotions are associated with wage increases, higher wage increases are associated with higher effort, and teachers increase effort in years leading up to promotion eligibility but reduce effort if they are repeatedly not promoted. Evaluation scores are positively associated with time spent on teaching and with student test scores, diminishing concerns that evaluations are manipulated.

JEL Codes: J31, J33, J45, M51

Key words: teacher incentives, promotions, China

*A previous version of this paper was titled Promotion Incentives and Teacher Effort in China. We thank Matthew Collin, Jishnu Das, Stefan Dercon, Eliana La Ferrara, Paul Glewwe, Clare Leaver, Lara Lebedinski, Derek Neal, Simon Quinn, Margaret Stevens, Francis Teal, Xuehui An, and seminar participants at Oxford, Bocconi, CSAE, CUHK, HKUST, IFPRI, ANU, Monash, UNSW, Fudan, and Peking University for helpful comments. This paper is produced as part of the project Actors, Markets, and Institutions in Developing Countries: A micro-empirical approach (AMID), a Marie Curie Initial Training Network (ITN) funded by the European Commission under its Seventh Framework Programme Contract Number 214705 PITN-GA-2008-214705.

[†]International Food Policy Research Institute (IFPRI) (N.Karachiwalla@cgiar.org).

[‡]Department of Economics, Division of Social Sciences, and Institute for Advanced Studies, HKUST, IZA, and CEPR (albertpark@ust.hk).

1 Introduction

Public servants comprise a large proportion of formal employment in low-income countries, but service delivery is often poor (World Bank, 2004). In many developing countries where teachers account for a large share of public sector workers, there are widespread complaints about high rates of teacher absenteeism and poor learning outcomes (Glewwe, 2002; Bruns et al., 2011). Similar concerns are voiced in developed countries about systems in which teachers enjoy high job security and pay is determined by education levels and seniority rather than by performance; these teacher characteristics are only weakly correlated with student learning (Podgursky and Springer, 2007). As a result, there has been great interest in reforms that incentivise teachers to exert greater effort, in particular through pay for performance schemes that link bonuses to student test scores (Muralidharan and Sundararaman, 2011; Neal and Schanzenbach, 2010).

China has taken a much different approach to incentivising teachers, and civil servants more generally, by establishing a sophisticated system of annual performance evaluations that inform promotion decisions (Ding and Lehrer, 2007b). In Chinese public schools, teacher absence is very rare, and student achievement is sometimes spectacular, as evidenced by students in Shanghai scoring top in the world in the Programme for International Student Assessment (PISA) exams (OECD, 2010). However, the design and performance of China's public service promotion system has received little attention or systematic study.

In this paper, we provide empirical evidence on how China's promotion system influences the effort of public employees. We test theoretical predictions of a tournament model of promotions using panel data on primary and middle school teachers in a province in western China, and find evidence that effort responds to promotion incentives in a manner remarkably consistent with theory. Specifically, we show that in Chinese schools, promotions are associated with substantial pay increases, effort is greater on average when the wage increases are larger, effort increases as the teacher gets closer to the year in which she is eligible for a promotion, teachers who are repeatedly passed over for promotions reduce effort substantially, and the larger the number of teachers competing for a promotion, the lower the effort of teachers at the extremes of the skill distribution. Although we focus attention on teachers, key features of the evaluation and promotion system are the same for all public employees in China. Thus our results may help shed light on the puzzle of how China has grown so rapidly despite having weak institutions for protecting property rights and guarding against corruption, which some argue threaten the sustainability of China's growth (Acemoglu et al., 2012).

Nearly all previous studies of promotion systems focus on the private sector. Of these, many study wage determination within companies to test whether pay increases substantially with promotion rather than reflecting current marginal productivity (Medoff and Abraham, 1980, 1981; Baker et al., 1994a,b). Only a few link promotion incentives to direct measures of effort or performance (Gibbs, 1995; Kwon, 2006; Campbell, 2008).

The performance of promotion systems may differ in the public sector in comparison to the private sector. Work output or productivity may be more difficult to quantify (Dixit, 2002), and public service jobs may involve multiple tasks or multiple principles to a greater degree than private sector jobs, making optimal incentives for specific tasks weaker (Holmstrom and Milgrom, 1991; Dixit, 1997; Rasul and Rogger, 2013). Teacher performance pay schemes have generally produced positive results on the outcomes they incentivize (i.e., test scores) (Lavy, 2003; Glewwe et al., 2010; Muralidharan and Sundararaman, 2011; Muralidharan, 2011; Barlevy and Neal, 2012); however, they have also been subject to teachers gaming the system (Jacob and Levitt, 2003; Figlio and Winicki, 2005; Figlio and Getzler, 2006; Neal, 2011), may also have dynamic effects on teacher effort (Macartney, 2014).

Unfortunately, there are almost no rigorous empirical studies of promotion systems in the public sector. Ashraf et al. (2014b) find that career incentives induce the selection of more career-oriented individuals into the health services sector in Zambia. Similar to the findings of Dal Bó et al. (2013); Ferraz and Finan (2009); Propper and Van Reenen (2010) such individuals are also more effective at delivering public services. Like most work on the private sector, many existing studies of public sector promotions investigate the determinants of wages or promotion without directly examining whether effort responds to incentives. Previous research has found that provincial and city officials in China are more likely to be promoted if they achieve high rates of economic growth (Li and Zhou, 2005; Landry, 2008) and that US mayors receive higher pay if they promote economic growth (Enikolopov, 2011). A study by Haeck and Verboven (2012) also examines promotion incentives in the education sector, showing that the wages of professors of a European university do not respond to the external labor market, and that those with better research and teaching performance are more likely to be promoted. These studies do not distinguish clearly whether the link between performance and promotions reflects selection or incentive effects. Ashraf et al. (2014a) is unique in examining incentive effects across the ability distribution.

The rest of the paper is organized as follows. The next section describes the promotion system for teachers and public servants in China. Section 3 introduces the data and presents descriptive information on how the teacher promotion system operates in practice. Section 4 presents a tournament model of promotions as incentives, which generates theoretical predictions to be tested empirically. Section 5 presents the empirical specification and identification strategy. Section 6 presents the estimation results, and Section 7 concludes.

2 China's Promotion System

In China, all civil servants compete with their colleagues for promotions based on extensive annual performance evaluations. In the education sector, teachers in primary schools can be assigned to four different ranks: intern, primary level 2, primary level 1 and primary high level. In middle school, teachers can be assigned to one of five ranks: intern, middle level 3, middle level 2, middle level 1 and middle high level. Primary level 2 is equivalent to middle level 3, primary level 1 is equivalent to middle level 2, and primary high level is equivalent to middle level 1 in terms of administrative status. Teachers are all hired as interns and then promoted.[1] Promotion to a higher rank is accompanied by an increase in salary, and the number of available spaces at higher ranks in each education district is limited.

There are specific rules regarding the years of service required before a teacher is eligible to apply for promotion to the next rank. In Gansu province, the study site, all teachers begin as interns in their first year, regardless of educational background, and can apply immediately for primary level 2 or middle level 3 in the second year of teaching. Promotion from intern to primary 2 or middle 3 levels is essentially automatic and do not really follow the bureaucratic rules outlined below. As such, we exclude them from most of the analysis. After obtaining these ranks, teachers with a middle school education can apply for primary level 1 or middle level

[1] This is one of the key features of internal labour markets, as defined by Doeringer and Piore (1971).

2 rank after 4 years, those with a vocational college degree can apply after 3 years, and those with a university degree after 1 year. Teachers who graduated from a middle school can apply for primary high or middle level 1 ranks after 10 years, with a vocational college degree after 7 years, and those graduating from a university after 5 years. To apply for middle high rank all teachers must have been at middle level 1 rank for at least five years, and teachers with middle school degrees also must have had at least 25 total years of service, while those with vocational college degrees must have 15 years of total service.[2]

Apart from needing to wait the requisite number of years before applying to the next rank, promotion applications also require the attainment of minimum scores on annual performance evaluations. Annual evaluations are conducted on a four point scale: excellent, good, pass, and fail. To be eligible for promotion, teachers must obtain either two 'good' annual performance evaluation scores or one 'excellent' evaluation score in the past five years. Other promotion requirements may vary by county or education district (typically comprising all of the schools in a township). For example, in some education districts teachers must publish an article in a teaching journal.[3] There are also teaching awards for teachers at the county, education district, or school levels which can influence promotion eligibility.

Each year the Education Bureau of each county determines the specific number of promotions available for each rank in each education district, and teachers are informed of the quota numbers by the district education officials. All teachers who have fulfilled the minimum conditions are then considered for rank promotions given the available spaces.

Annual teacher performance evaluations assess multiple dimensions of teachers' performance, grouped into four categories: 1) student test scores, 2) work ethic and attitude, 3) preparation (lesson plans, homework, reporting etc.) and 4) teacher attendance. A point system is used, and although there is variation, student test scores[4] generally have the highest weight (34.2% on average), followed by preparation (29.5%), attitude (23.2%), and attendance (13.2%). Weights for each category and the specific indicators and points used to evaluate performance within each category are determined by a district committee typically comprised of principals, vice-principals, and outstanding teachers from schools in the district. Interviews with teachers made

[2]Teachers with PhDs can apply after only one year.

[3]The requirements for this also vary, with some counties simply requiring a summary of the teacher's teaching experience, and some counties ranking publications by the quality of the journal.

[4]In some townships, mean test scores are used, and in other townships, value added test scores are used.

clear that most do not have a precise idea of what the specific weights are, but they do know the components. The committee draws upon information from assessments based on direct classroom observation, assessments by other teachers in the school (peer review), questionnaires from students, data on student test scores, and principal reports of the teacher's attendance, preparation, and attitude. Once final scores are agreed upon, there are fixed percentages of teachers who recieve 'excellent' and 'good' evaluation scores. The data shows that approximately the top 10-15% of teachers *in each district* receive 'excellent' evaluation scores and the next 30-40% receive 'good' evaluation scores, and this is consistent across ranks. The names of the recipients of excellent and good evaluation scores are announced at an annual meeting of all teachers. Since there is an explicit quantitative formula and the results are announced publicly, the ability of committees to manipulate evaluation scores is limited; however, this practice will vary by district. The district committee in charge of evaluations also makes recommendations on who should be promoted each year subject to final approval by the county Education Bureau, based on another point system that incorporates the history of annual evaluation scores, publications, awards, and other district-specific criteria.

Principals also are evaluated annually using a point system and receive promotions and salary increases based on the performance of their school. A committee appointed by the county Education Bureau visits each school every year and observes the school's facilities and grounds, collects test score data, surveys teachers about the principal's management of the school, looks through the school accounts, etc. This process helps make the system of promotions incentive compatible, avoiding to a great degree the problems of influence activities for which promotions are often criticised (see Fairburn and Malcomson, 2001).

3 The Gansu Survey of Children and Families

The Gansu Survey of Children and Families (GSCF) is a longitudinal study of rural children in Gansu province in northwest China. The main survey sampled 2,000 children aged 9 to 12 in the year 2000 living in 100 villages in 42 townships in 20 counties. In order to better understand the supply side of the educational system, all principals and teachers in the main primary and middle schools attended by the sample children were also surveyed, providing detailed information on the background of principals and teachers, their time allocation, and their remuneration. Generally speaking, in rural counties in Gansu there is one primary school in each village and one middle school in each township. Each sample child can be linked to his or her homeroom teacher. A second and third wave of the GSCF followed the same children in 2004 (at age 13 to 16) and 2007 (at age 16 to 19). In 2007, individual and household questionnaires, as well as achievement tests, were also administered to a new cohort of 1,500 children aged 9 to 15.

This study uses a panel dataset constructed for teachers spanning the years 2003 to 2006, based on surveys of 2,100 teachers collected during the third wave of the GSCF in 2007. It was only in the third wave that retrospective data was collected on past teacher evaluation scores, as well as full promotion histories. The average sampling rate of teachers in the schools is 85% and only 23% of the schools have a sampling rate lower than 25%.[5]

Table 1 provides summary statistics on teachers in the sample, including breakdowns by teacher rank in 2007. Among the teachers, 38% are female. Teachers in higher ranks tend to be male, older, more experienced, better educated, and more likely to have received teaching awards or published papers in teaching journals.[6] Teachers in higher ranks also performed slightly better on their teacher placement tests (on education, psychology, and language) taken at the beginning of their teaching career, which we take as an indicator of ability prior to having any teaching experience. Of course, some of these differences in performance across ranks may be related to differences in years of experience. Teachers at lower rank levels spend more hours per week on grading, lesson plans, and with students, which could reflect greater motivation or incentives, or a learning effect if more experienced teachers use their time more efficiently. Teachers with higher ranks also teach slightly fewer classes. Interestingly, teachers with the highest rank (middle high) work substantially fewer hours. As expected, salaries increase substantially with rank. The biggest jumps are from primary 2 to primary 1, followed by middle 3 to middle 2, and then middle 1 to middle high. As the final row of Table 1 shows, the average salary increases are 32.4%, 25.1%, and 21.5%, respectively.

Table 2 provides a breakdown of evaluation scores by rank for the years 2003-2006. For

[5]The sampling rate is calculated from the total number of teachers in the school as reported in the Principal's questionnaire. The rate is: number of teachers interviewed / total number of teachers reported by the principal.

[6]These characteristics are consistent with the findings of Ding and Lehrer (2007b) describing teachers in much richer Jiangsu Province.

most teachers, four years of data are recorded, so the total number of teacher-year observations of evaluation scores is 8,713. Not many teachers receive an evaluation score of 'fail', and the majority of the evaluation scores are 'pass'. There is also evidence that strict limits are enforced on the share of teachers who can obtain good evaluations, as the proportions of teachers receiving each type of evaluation score are quite similar across ranks.

Tables 3 and 4 contain the results from wage regressions for primary school and middle school teachers, respectively. The first column is a regression of the log of monthly wages on rank level dummy variables; the second column adds controls for teaching experience; and the final column includes dummy variables for the highest level of education completed, as well as the performance evaluation score for the most recent year (2006). All regressions include county fixed effects and standard errors clustered at the county level. The omitted rank levels are primary 2 for the primary school teacher regressions, and middle rank 3 for the middle school teacher regressions. Rank level premiums are significant and increasing at each higher rank. At the primary level, the increase from primary 2 to primary 1 is 27% and to primary high is an additional 16%. For middle school teachers, the increase between levels is 20%. Controlling for individual characteristics reduces these premiums, but they remain high, particularly at the top ranks.[7] These results show that promotions are rewarded by salary increases independently of experience, education, and ability, providing prima facia evidence that promotions are used to incentivise effort. Similar results form the main findings of earlier empirical studies of promotion systems in the private sector.

The wage regressions also reveal that in addition to increasing with each higher rank, wages rise modestly with experience within each rank. We note that because county Education Bureaus have a local monopoly (there are almost no private schools in these areas) and determine assignments of teachers to schools, there is very limited mobility of teachers across schools, and even less across counties because county Education Bureaus can administratively block teachers from making such a move. Teachers are only moved for specific reasons, such as a change in work location of a spouse, or because they have subject-specific skills that are needed in another school or county. Salaries are very similar across public schools within the same

[7]Variation in wages is mainly driven by rank and experience. There is less variation across counties. When the results of Tables 3 and 4 are compared to regressions that do not include county fixed effects, there is little change in the $R^{2\prime}s$.

counties, providing little incentive to switch schools within the county to pursue higher pay. Han (2013) notes that beginning in 2001, Gansu began to centralize teacher allocation across schools. She finds, however, that differences in wages and other benefits are not significantly different across counties that did and did not centralize early.[8] Only 20 teachers (0.85%) in the sample reported switching counties between 2003 and 2007.[9] China's residential registration system links access to public goods and services to one's place of official registration, which also inhibits labor mobility (Chan et al., 2008).

Figure 1 shows how long it takes to be promoted to each of the rank levels. Figure 1 displays Kaplan-Meier survival functions for how long a teacher 'survives' at each rank level before being promoted to the next rank. Most teachers who get promoted from primary level 2 to primary level 1 do so within 4 to 5 years. Promotion is nearly automatic; practically everyone is promoted at some point. For promotions from primary level 1 to primary high level, the distribution of the number of years until promotion is far more spread out. There tends to be a steady stream of promotions, but just under 1/3 of primary level 1 teachers are never promoted to primary high. At middle level 3 rank, promotion also is almost automatic, with every teacher promoted within 9 years. The promotions from middle level 2 to middle level 1 look similar to those from primary level 1 to primary high. They are more spread out, but the majority occur between 5 and 10 years. Approximately 25% of those at middle level 2 are not promoted to middle level 1. Very few people (less than 25%) are promoted from middle level 1 to middle high level in this sample, and the promotions that do occur take considerably longer (at least nine years). The differences in the promotion rates and timing across different ranks provide a good opportunity to study differences in effort due to differing promotion rates, as well as to study how effort evolves over time.

The next section presents a model of promotions that allows us to develop an estimating equation to test the extent to which the promotion system rewards and motivates effort.

[8]In related work, Han et al. (2014) find that schools in Gansu that centralized teacher deployment faster from 2001 to 2003 performed more poorly in terms of test scores and teacher work hours. Our data are from 2007 by which time the centralization fo teacher deployment was long completed.

[9]All the results are robust to dropping these 20 teachers (results not reported).

4 A Model of Promotions as Incentives

We present a model of promotion incentives based largely on that of Gibbs (1989) and Gibbs (1995), which yields testable predictions that are then taken to the data in the empirical analysis. First, consider a one-period tournament model in which n teachers in a district compete for a fixed number of promotion slots, k. Wage increases associated with promotion are taken to be exogenous. Teachers differ in skill, s, which is not directly observable by principals. Teachers know their own skill but not that of others. We assume that the skill of each teacher is drawn from a uniform distribution centered around zero, which is common knowledge. Teachers exert effort, e, in their job (capturing both hours and intensity of effort), which is costly and gives them disutility $C(e)$, and is assumed to increase in e at an increasing rate ($C' > 0$ and $C'' > 0$). Performance is measured by $q_i = s_i + e_i + v_i$, where q_i is output and $v_i = \epsilon_i + \mu$ is an error term, which comprises individual error (or 'luck') and a common error. Output does not include simply test scores; it also includes overall learning and student attainment. We assume that v_i has a symmetric distribution around zero and $E(v_i) = E(\epsilon_i) = E(\mu) = 0$. Let $F(\epsilon)$ be the CDF and $f(\epsilon)$ be the PDF of the individual error term. We further assume that $d^2q/de^2 = 0$ and $d^2q/dsde = 0$, which implies that there are no scale economies in effort and no productive interactions between effort and skill.

Each teacher's probability of promotion depends on her skill and effort, and the skill and effort of others. Thus, a teacher's probability of promotion can be written as $p = p(e, s, \underline{e}, \underline{s})$ where $\underline{e}$ and $\underline{s}$ are vectors of other teachers' efforts and skills, respectively. The county government offers a reward for promotion, ΔEV, which is the change in expected lifetime utility gained from winning a promotion relative to not being promoted. This change mainly reflects the greater lifetime income associated with the wage increase that comes with promotion (for example, $(W_2 - W_1)$ if wages for rank 2 are higher than for rank 1) but could also reflect non-pecuniary benefits and the option value of possible future promotions.

The teacher chooses effort to maximise the expected reward minus the disutility of effort:

$$\max_{e} \{p(e, s, \underline{e}, \underline{s})\Delta EV - C(e)\} \tag{1}$$

The first order condition sets the marginal cost of effort equal to the gain in utility, times the marginal effect of effort on the probability of promotion:

$$C'(e) = \frac{dp(e, s, \underline{e}, \underline{s})}{de} \Delta EV \tag{2}$$

Define the derivative $\frac{dp}{de}$ the marginal probability of effort (MPE). The second order condition is satisfied as long as the slope of the probability function does not increase too quickly in effort:

$$\frac{d^2p}{de^2} \Delta EV - C''(e^*) \leq 0 \tag{3}$$

To win a promotion, a teacher must place k^{th} or higher, in other words beat at least $n - k$ workers in pair-wise comparisons of performance. A teacher treats the effort of a competitor as a function of the random variable that is the competitor's skill, and treats all competitors as ex-ante identical in deriving the MPE. Given CDF $R(q_j)$ and PDF $r(q_j)$ over a competitor's performance, the probability that teacher i beats teacher g is $pr(q_i > q_g) = pr(e_i + s_i + \epsilon_i + \mu > e_g + s_g + \epsilon_g + \mu) = pr(e_g + s_g + \epsilon_g < e_i + s_i + \epsilon_i) = R(e_i + s_i + \epsilon_i)$. The probability that teacher i beats any opponent is $R(e_i + s_i + \epsilon_i)$ and the probability that teacher i loses to any opponent is $1 - R(e_i + s_i + \epsilon_i)$.

To finish exactly j^{th} from the top, a teacher must beat $n - j$ opponents and lose to $j - 1$ opponents. The number of ways to choose $n - j$ elements from a collection of $n - 1$ elements is $\binom{n-1}{n-j} = \frac{(n-1)!}{(n-j)!(j-1)!}$ and the probability of promotion conditional on ϵ_i is the sum of the conditional probabilities of placing first through k^{th}. Integrating out ϵ_i gives the unconditional probability of promotion:

$$p(e_i, s_i, \underline{e}^*, \underline{s}) = \sum_{j=1}^{k} \binom{n-1}{n-j} \int R(e_i + s_i + \epsilon_i)^{n-j} [1 - R(e_i + s_i + \epsilon_i)]^{j-1} f(\epsilon) d\epsilon_i \tag{4}$$

Here, the only stochastic part of the teacher's own performance over which he or she must integrate is ϵ, and the distribution of skills and effort of competitors influence the expected probability of beating a given opponent (R). If the overall promotion rate is equal to either zero or one ($k = 0$ or n) so that either no-one or everyone is promoted, then $p(e^*, s, \underline{e}^*, \underline{s})$ is constant, the MPE is zero, and no teacher has an incentive to exert effort.

To solve for the MPE we need to differentiate with respect to e. Rearranging terms, and defining the change of variables $\Psi \equiv R(e + s + \epsilon)$[10], we get the following:

$$\frac{dp}{de} = \int (n-k)\binom{n-1}{n-k}\Psi^{n-k-1}(1-\Psi)^{k-1} f(R^{-1}(\Psi) - s - e)d\Psi \tag{5}$$

.

Here, the subscript i has been dropped for notational convenience. Optimal effort is determined by the Nash equilibrium solution in which each teacher exerts effort anticipating the skill and effort choices of his or her competitors (which affects R).

Consider the case in which opponents' effort is fixed or when variations in optimal effort across skill types are small relative to variations in s and ϵ. In this case, for a teacher with a given skill level, q_j will be distributed symmetrically reflecting the random distribution of luck. Apart from $f(R^{-1}(\Psi) - s - e)$, (5) has the form of a Beta density function $\beta(k, n-k)$. Some relevant properties of the Beta densities are as follows. First, for $k, n-k > 1$, the Beta desnsity is uni-modal with mean $(n-k)/n = 1 - p^*$. Second, $\beta(k, n-k)$ is symmetric about one half if and only if $k = n/2$ (when half of the teachers are promoted). $\beta(k, n-k)$ is skewed to the right if $k > n/2$, is skewed to the left if $k < n/2$, and becomes increasingly skewed as k/n deviates from $1/2$ in either direction. Finally, holding k/n fixed, as n (and k) approaches infinity, $\beta(k, n-k)$ approaches a distribution with all of its mass at $1 - p^*$.

To develop intuition, assume for a moment that all teachers are homogeneous in skill. In Figure 2 the bottom half of Panels A, B and C show the (re-scaled) distribution of luck draws, which is the same regardless of the promotion rate. Middle values of luck are most likely; very high or very low draws of luck are less likely. The MPE is a weighted sum of the Beta density values (from the top panels), where the weights are the f (luck) distributions plotted in the bottom panels. Panels A, B, and C then represent the MPE for every draw of luck when $p^* = 1/2$, $p^* < 1/2$, and $p^* > 1/2$, respectively. The weighted average is highest when $p^* = 1/2$ since in this case high Beta density values are given greater weight and low Beta density values are given less weight. This means promotion incentives are highest when the promotion rate is 1/2. Intuitively, all teachers have an equal chance (50%) of promotion ex-ante, so all teachers are on the margin of being promoted versus not being promoted and small amounts of luck

[10]See Appendix C for full derivation.

can decide who wins the promotions. In this case, effort makes the biggest difference to one's probability of promotion. If the promotion rate were greater than 1/2, a teacher would need a low draw of luck in order not to be promoted. Incentives are lower here, since more than half of teachers will get promoted and low values of luck are unlikely. In this case, teachers do not have to work as hard to be promoted. Similarly, if the promotion rate were less than 1/2, a teacher would need a high draw of luck in order to be promoted. Incentives are lower here as well, since few teachers will be promoted and high draws of luck are unlikely.

Now consider the case of teacher heterogeneity with respect to skill, again holding the effort of competitors constant or assuming variations in optimal effort are sufficiently small. For a teacher with average skill, incentives are highest when the promotion rate is 1/2, as they will be on the margin of being promoted versus not being promoted. But at $p^* = 1/2$, high skill teachers do not need to work as hard, as they would require a very low draw of luck in order not to be promoted, which is unlikely. Figure 3 shows that $f(R^{-1}(\Psi) - s - e)$ will shift to the right over [0,1] for high skill teachers. Here, low Beta density values are given high weight. Skill has the effect of decreasing the luck required to win. So incentives are low for high skilled teachers. Similarly, $f(.)$ will shift to the left for low skill teachers, who would need a very high (and unlikely) draw of luck in order to be promoted, and effort incentives are low for these teachers as well.

Define a marginal skill teacher as one whose skill shifts $f(.)$ so that it peaks in the same region as the beta density in (5). The marginal skill level depends on the promotion rate, p^*, since the mode of $\beta(k, n-k) = 1-p^*$. This is the teacher for whom effort incentives are highest; it is exactly the point at which the highest Beta density values are given the highest weight. Intuitively, if for example 1/3 of teachers are promoted, the teacher with the highest incentives is the one who believes she is at the 2/3 skill percentile (from the bottom). Such a teacher is on the margin between being promoted and not being promoted, and a small amount of luck can determine whether she wins the promotion or not, so effort will make the most difference to her probability of promotion. Incentives decrease monotonically with distance from the marginal skill percentile.

Now consider what happens to the marginal probability of effort when teachers account for the effort choices of competitors. Since the effort of competitors increases as a teacher's skill percentile gets closer to one minus the promotion rate, the marginal probability of effort curve will become more tightly distributed around one minus the promotion rate (given effort choices of competitors it is harder to get promoted with skill and luck draws farther from one minus the promotion rate). This further reinforces the result that effort incentives decline with the distance between a teacher's skill percentile and one minus the promotion rate.

Effort incentives also depend on the number of teachers competing for promotion. When n increases but p^* stays the same, Gibbs (1989) finds that incentives increase when p^* is between 1/3 and 2/3 in the case of homogeneous teachers. With more competitors, the game becomes more precise in that the variance of the threshold score required to win is reduced. The Beta density becomes taller at $1 - p^*$. In the case of heterogeneous skill, for teachers with skill at or near the marginal skill percentile, incentives increase since it is now more worthwhile to exert effort to beat luck. However, for those further from the margin, incentives decrease. Figure 4 illustrates this point, showing that effort incentives increase for teachers whose skill percentile falls between the arrows, and decreases for those whose skill percentile falls outside of them.[11] Teacher A's skill percentile is close to the marginal skill percentile, but teacher B's is further away.

Thus far, we have modeled a one-period tournament. In reality, evaluations occur every year while promotions occur after multiple years of teaching, and failure to be promoted in one year does not preclude promotion in future years. Although many key aspects of how promotions influence incentives are captured by the one-period model, thinking about the model in terms of multiple periods enables further predictions about how effort responds to the timing of promotion incentives and updated beliefs about a teacher's relative skill rank. We focus first on how a specific rule in China's promotion system influences effort incentives over time. As described earlier, to be eligible for promotion a teacher must have completed a certain number of years of service (depending on education and experience) and obtained one 'excellent' or two 'good' evaluations within the past five years. Here, we provide a simplified multi-period model to capture how these rules affect optimal effort incentives.

Assume that teachers have careers lasting T periods. They are eligible for promotion after serving X years. Teachers can be promoted in any year t after they become eligible ($t > X$). As

[11] If the promotion system were based on a standard, the number of teachers competing would not matter.

noted earlier, observed teacher performance in each year t is equal to the sum of skill, effort, and luck: $q_t = s + e_t + v_t$. The expected probability of promotion (p_t) in any given year t is based on teacher performance in the previous five years ($p_t(q_{t-1}, q_{t-2}, q_{t-3}, q_{t-4}, q_{t-5})$). Two simplifying assumptions are made: (1) that the impact of effort in one year on the probability of future promotion is independent of the impact of effort in other years, that is $d^2p_t/de_{t-i}de_{t-j} = 0$ for all i and j, and (2) that performance of effort in any of the five years before promotion has an equal impact on the probability of promotion ($dp_t/de_{t-1} = dp_t/de_{t-2} = ... = dp_t/de_{t-5}$).[12]

Now consider a teacher's optimal effort decision in year t. We can normalise the per-period utility from wages before promotion to equal zero, and define $U_h > 0$ to be the per-period utility from wages after promotion. In any given year j, the elements of lifetime expected discounted utility that are affected by effort in that year can be expressed as:

$$EV_j = -c(e_j) + Ep_{j+1} \sum_{t=j+1}^{T} \beta^{t-j} U_h + (1 - Ep_{j+1})[Ep_{j+2} \sum_{t=j+2}^{T} \beta^{t-j} U_h + (1 - Ep_{j+2})\{Ep_{j+3} \sum_{t=j+3}^{T} \beta^{t-j} U_h + (1 - Ep_{j+3})(Ep_{j+4} \sum_{t=j+4}^{T} \beta^{t-j} U_h + (1 - Ep_{j+4}) Ep_{j+5} \sum_{t=j+5}^{T} \beta^{t-j} U_h)\}] \quad (6)$$

Here, $c(e_j)$ is the disutility of effort in year j, β is the discount rate, and Ep_t denotes the expected promotion probability given the distribution of draws of luck, where $p_t(s, e, \underline{e}, \underline{s})$ is defined as before. Current effort can alter the probability of being promoted in the subsequent five years, and once promoted the teacher enjoys higher utility until period T. Note that beyond year $j + 5$, the expected probability of promotion is not influenced by current year effort.

As thus formulated, maximisation of (6) with respect to e_j defines optimal effort by the teacher. If a teacher's beliefs about s are fixed, then the only thing that changes over time is that T becomes nearer so that there are fewer future years of higher wages to enjoy, reducing the expected benefits of promotion, and leading to lower effort over time.

Given the simplifying assumptions, effort in the current year has an identical marginal effect on expected promotion probability in the subsequent five years unless the teacher is ineligible

[12]These assumptions are made for tractability. As long as effort and performance in each of the previous five years increases the probability of promotion in years of promotion eligibility, the qualitative results are essentially the same.

for promotion in those years ($t \leq X$). In those years, the expected probability of promotion is zero, as is the marginal probability of effort in all previous years. Thus, as long as $t \leq X - 5$, there is no incentive to exert effort because in all of the subsequent five years, the teacher is ineligible for promotion. It is easy to show that in the years from $X - 4$ to X, the marginal probability of effort increases because with each additional year of service the number of future eligible years that can be influenced by current effort increases by one year, and the first year of possible promotion is closer to the present increasing the discounted payoff from promotion (see proof in Appendix D). Figure 5 illustrates this point. In year $t = X - 4$ for example, there is one year in which current effort affects probability of promotion and the teacher is eligible for promotion ($t = X + 1$). In year $t = X - 3$ there are two such years ($t = X + 1$ and $t = X + 2$). In year X and all subsequent years, the teacher is eligible for promotion in each of the subsequent five years (there are always five 'eligible years'), so there is no longer an incremental increase in effort incentives over time associated with changes in the promotion eligibility window.

Next, we consider how being repeatedly passed over for promotion is expected to affect effort incentives. The simple intuition is that continued failure should discourage teachers, leading to lower effort. Assume that in year 1, teachers begin with a prior belief (s_1) about their skill, while their true skill is s. This belief can change over time (s_t) as teachers update their beliefs based on new information. As shown earlier, the marginal probability of effort is maximised when $s_t = 1 - p^*$, or when one's relative skill position coincides precisely with the share of teachers promoted. As the absolute value of $s_t - (1 - p^*)$ increases, the marginal probability of effort falls. Assume that if a teacher initially (in period 1) believes that her skill level is s_1, new information about her performance enables the teacher to update beliefs about her skill level as well as the skill level of other teachers.

We focus on one key piece of information about relative performance that is observable to both the teacher and the econometrician – whether the teacher was promoted or not in the past year. Because the probability of promotion is a positive function of both skill and effort in the previous year, the lack of promotion in a year in which the teacher is eligible for promotion provides an unambiguously negative signal to the teacher about her relative skill level. Using Bayesian updating, if a teacher is passed over for promotion in period t, the teacher will adjust downward her estimate of her relative skill level. In other words, $s_t < s_{t-1}$ if the teacher is not promoted in period $t > X$.

Of course, a number of factors influence the extent to which the teacher believes the failure to be promoted is informative about her skill level. For example, if very few teachers can be promoted because of few quota positions available for higher ranks (the implied $1 - p^*$ would then be very high) or if the teacher knows that some other teachers are better performers and are likely to be promoted before her (meaning that she believes her skill percentile is further from $1 - p^*$ than that of her competitors), then it may take repeated promotion failures before she substantially revises downward her assessment of her relative skill level. However, as additional years pass without promotion, the teacher's estimate of her relative skill level should fall, further increasing the gap between estimated relative skill s_t and $1-p^*$. From earlier, we know that the larger this gap, the lower the marginal probability of effort. Thus, if the failure to be promoted is revealing information to teachers about their skill level, we would expect a teacher's effort to decline as she continues to be passed over for promotion year after year.[13]

Given that teachers compete over multiple years for promotion, the annual evaluation scores can be viewed as interim feedback that, similar to being passed over for promotion each year, enables teachers to update their subjective beliefs about their skill and that of others. There is an emerging literature that examines how feedback influences the outcomes of dynamic tournaments. Only a few studies consider heterogeneous contestants. Ederer (2010) shows that in a model with two contestants with differing ability, interim feedback has the effect of lowering effort incentives for both high and low ability types. The effect is more pronounced and equilibrium effort is lower when differences in ability are greater. When differences in ability are small, however, equilibrium effort is higher since both contestants are close to the margin.[14] Thus, interim information may help teachers learn about their skill, enhancing the validity of the model's predictions assuming that individuals are informed about their own skill level. However, in our context the informational content of annual evaluation scores remains limited in important respects.[15]

[13]The declining effort is the result of the promotions not being 'up or out' promotions.

[14]Hansen (2012) also finds that the optimal disclosure policy of a firm is to reveal information in the middle of the distribution, but not in the tails.

[15]Despite receiving interim feedback from annual evaluations, teachers are still unlikely to have very complete information about their skill level relative to their competitors. They often do not know the full history of evaluation scores and other credentials of other teachers, especially those who teach in other schools. They also do not know the record of other teachers on specific requirements for promotion other than annual evaluation

Unfortunately, it is beyond the scope of this paper to fully model or test the dynamic impact of interim feedback in the form of annual evaluation scores on subsequent effort. Given the promotion rules used in Chinese schools, it would be important to know the full history of evaluation scores of all teachers in a given rank in an education district to understand how they would influence subsequent effort. This data is not available. A practical concern is that if we were to regress annual evaluation scores on lagged scores, it would be difficult to disentangle incentive effects from mean reversion or serial correlation, especially given the short period of data available.

In summary, the tournament model of teacher promotions has yielded the following testable predictions:

1. Teachers will exert effort in order to win promotions;

2. A greater increase in wages corresponds to higher average effort incentives;

3. The closer is p^* to $1/2$, the higher is average effort;

4. Incentives are highest for teachers who believe they are on the margin of being promoted versus not being promoted, or whose skill percentile is equal to $1-p^*$. Incentives fall with the absolute value of the difference between teachers' skill percentile and $1-p^*$;

5. When n increases but p^* stays the same, increasing the number of teachers increases incentives for those with skill percentile near the margin, and decreases incentives for those further away;

6. Given China's specific promotion elibility rules, effort incentives are expected to be zero five or more years before a teacher becomes eligible for promotion (when $t \leq X-5$ where X is the first year of promotion eligibility) and increase over the five years preceding promotion eligibility (from $t = X-4$ to $t = X$);

7. If a teacher continues to be passed over for promotion when she is eligible for promotion, then her effort will decline (many years after $t = X$); and,

scores (i.e., publication in teaching journals, teaching awards, etc.). Importantly, teachers do not know how many teachers in their rank and district will be promoted each year, or in many cases even how many teachers are eligible for promotion (especially teachers in other schools). Even if a teacher has herself received either two good evaluation scores or one excellent score, she will still be unsure of her promotion probability, since that is a necessary but not sufficient condition for promotion.

8. After the highest promotion, effort incentives no longer exist ($p^* = 0$).

5 Empirical Framework

This section describes the empirical framework used to test the theoretical model's predictions. Our main focus is to estimate whether the determinants of teacher effort are the same as predicted by theory. In addition, we verify that effort is higher the larger are wage increases, and that better evaluation scores are in fact rewarded by a higher probability of promotion.

From the FOC (equation (2)) we see that effort depends on the gain in expected utility from promotion, in particular the wage gain that a teacher would receive if promoted, and the marginal effect of effort on promotion probability (MPE). The MPE depends on the number of competitors (n), the promotion rate ($p^* = k/n$), skill (s), and time to promotion eligibility. Effort incentives are expected to be greater the higher the wage increase from promotion, the closer is $1-p^*$ to the teacher's skill percentile, the larger the pool of competitors (if the teacher's perceived skill percentile is close to $1-p^*$), and the closer is t to the time a teacher becomes eligible to apply for a promotion:

$$\begin{aligned} &e = f(\Delta EV, MPE), MPE = h(n, p^*, s, t) \\ &e = e(\Delta EV, n, p^*, s, t) \end{aligned} \tag{7}$$

Equation (7) is estimated using annual performance evaluation scores as a measure of effort. This is justified by controlling for the teacher's skill (which also influences performance) as measured by education, experience, and ability, and in some specifications individual fixed effects. We estimate two regression specifications for (7)–one that controls for county fixed effects and another that controls for individual fixed effects. The first enables us to test the impact of wage increases, promotion rates, and individual time-invariant attributes such as ability on effort. The second enables us to control for individual heterogeneity in ability and motivation to better identify the impact on effort of time relative to promotion eligibility and changes in tournament conditions over time. The first equation is specified as follows:

$$ev_{irct} = \zeta_0 + \zeta_1 \Delta w_{rc} + \zeta_2 p^*_{rc} + \zeta_3 a_{ir\tau t} + \zeta_4 n_{r\tau t} + \zeta_5 X_{irct} + \gamma_c + \upsilon_{irct} \tag{8}$$

Here, ev_{irct} is the evaluation score of teacher i at rank r in county c in year t, Δw_{rc} is the increase in log wages if promoted to the next rank in the county, p^*_{rc} is the promotion rate for those with rank r in county c (we use the average historical rate at that rank in a county), $n_{r\tau t}$[16] is the number of teachers in the same rank in the same education district (τ), a_{irct} is the ability of teacher i, X_{irct} are controls, γ_C captures county-level unobserved, time-invariant factors, and υ_{irct} is an error term which includes omitted individual and district characteristics.[17]

Since wages and promotion rates are set at the county level, we may be worried that there are unobserved factors at the county level that affect the way in which wage rates and promotion rates are chosen. Including county fixed effects sweeps away potential bias from time-invariant county-level unobservables. We can still identify the impact of wage changes and promotion rates by exploiting the variation in these variables across ranks within the same county.[18]

We expect that ζ_1 will be positive and significant (Prediction 2). This would tell us that promotion contests with higher wage gaps elicit higher levels of effort. We include the promotion rates as the proportion of teachers in a county that historically are promoted at that rank within a 'reasonable' number of years[19], and split this into quintiles. The average promotion rate in the quintiles is 11%, 37%, 52%, 68%, and 89% in the first to fifth quintiles, respectively. We leave the third quintile as the omitted category, so we expect the coefficients on the remaining promotion rate quintile dummy variables to be negative, and the coefficients for the first and fifth quintiles to be the most negative. This will test Prediction 3.

As a proxy for ability we use an 'ability index' constructed using the teachers' scores on three tests (in education, language, and psychology) taken when teachers apply for their first teaching job before they begin teaching. Scores on these tests are included by first standardizing the scores by the year in which they were taken, and then for each year of the data (2003-2006) the

[16] The contests are between teachers in the same district, but the wages and promotion rates are set at the county level.

[17] We treat the evaluation score as a continuous measure, rather than estimating an ordered probit model. This is because ordered probit cannot be estimated with teacher fixed effects, and because probit regressions of the same form where the outcomes are a dummy variable for receiving an excellent evaluation score, or for receiving an excellent or good evaluation score, yield consistent results.

[18] A robustness check is also done that includes county level observables such as average age, proportion female, and average levels of education of teachers to the regression to account for factors that change over time. This did not affect the results.

[19] We first calculate the number of years it took for each teacher to be promoted from one rank to another (we have each teacher's entire promotion history). A 'reasonable' number of years is defined as the number of years within which, for each rank and education level overall, half of the teachers were promoted. Then for each rank and county we calculate the proportion of teachers that were promoted within that number of years. We only include promotions before 2003, so this is a historical promotion rate in the county.

following measure of relative skill is calculated: $[S - ave(S)]/ave(S)$ where S is the standardized score on the tests for a teacher, and $ave(S)$ is the average of the scores of the teachers in the same rank in the same education district. An index was constructed using Principal Components Analysis (PCA).[20] Because the model's predictions are on how far ability is from the promotion rate (Prediction 4), we calculate dummy variables for whether the teacher belongs to the top or bottom 10% of the ability distribution relative to other teachers in the same rank, as we expect these percentiles to be very far from the promotion rate.

To measure n, we use the (log of) number of teachers in the same rank in the district.[21] Since the prediction on the effect of the number of teachers competing differs for those at the extremes of the skill distribution (Prediction 5), interactions between the number of teachers and ability in the top and bottom 10% of the skill distribution are also included in addition to n. In the X_{irct} controls we include: the teacher's age, a dummy variable for female, and education dummy variables for having completed high school, vocational college, or university.

We also estimate a second equation for the determinants of evaluation scores that controls for individual teacher fixed effects. This specification is estimable thanks to the availability of panel data, and has the great advantage of controlling for time-invariant individual-level heterogeneity in skill or motivation. The estimating equation is the following:

$$ev_{it} = \delta_1 \boldsymbol{D_{it}} + \delta_2 a_{10,it} + \delta_3 a_{90,it} + \delta_4 n_{it} + \delta_5 n_{it} * a_{10,it} + \delta_6 n_{it} * a_{90,it} + \phi_i + v_{it}. \qquad (9)$$

Here, $\boldsymbol{D_{it}}$ is a set of dummy variables for years before and after eligibility for promotion, $a_{10,it}$ is a dummy variable for whether the teacher falls into the bottom 10% of the ability distribution (10th percentile), and $a_{90,it}$ is a dummy variable for whether the teacher falls into the top 10% (90th percentile) of the ability distribution in that rank in the district, n_{it} is the (log) of the number of teachers in the rank in that district, ϕ_i is the individual teacher fixed

[20] Another index was also constructed that standardized each of the variables and then summed them. This did not affect the results.

[21] The number of teachers in each rank is adjusted by the undersampling rate at the school. This is calculated using the total number of teachers in the school as reported by the school principal. The principals did not report on the total number of teachers by rank, so the total number reported was compared to the number of teachers interviewed in the school to calculate the sampling rate and the undersampling rate of the school. The number of teachers in each rank was then adjusted as follows: number interviewed * (1 + undersampling rate). This assumes that that there were no systematic differences in the number of teachers that were not interviewed by rank, and the adjustment was applied in the same way for each rank.

effect, and v_{it} is the error term. The change in wages and the promotion rate drop out with the inclusion of individual fixed effects. Note that by controlling for teacher fixed effects, the coefficients on the ability variables and n (plus their interaction) are identified from changes in the relative of ability of teachers and number of teachers over time as the pool of competitors changes (due to retirement and promotions). The inclusion of n interacted with the ability dummy variables provide another test of Prediction 5. We expect the coefficients on these interaction terms to be negative ($\delta_5 < 0$ and $\delta_6 < 0$).

In $\boldsymbol{D_{it}}$ we include dummy variables separately for each year from $t = X - 4$ to $t \geq X + 15$. With the reference period being $t \leq X - 5$, the period five years and more before promotion eligibility, we expect that the coefficients on the time dummy variables for $t = X - 4$ to $t = X$ should be increasing. After many years of promotion eligibility, we expect that the coefficients will begin declining. Thus, we predict an inverted-U shape in performance with respect to years in a given rank. These dummy variables will test Prediction 6 and Prediction 7.

In the regression analysis, we focus on three rank levels for the effort regressions: primary 1, middle 2, and middle 1. This is because promotions are almost automatic for those with lower ranks (primary 2 or middle 3) and there are no further promotions after primary high and middle high. Thus, the teachers with primary 1, middle 2 and middle 1 rank levels have the clearest promotion incentives. However, to see the incentive effects for teachers with the highest rank levels (Prediction 8), we run similar regressions for teachers with rank levels of primary high or middle high, but instead of pre- and post-promotion eligibility time dummy variables, we use dummy variables for the number of years into the promotion (from two years after promotion to primary high or middle high rank levels to fifteen or more years after promotion).

After examining the determinants of teacher effort, we next turn to the determinants of being promoted. We test whether teachers' probability of promotion increases with teacher effort, skill, and the promotion rate as predicted by the theory. That is,

$$p_i = p(e_i^*, p^*, s_i) \tag{10}$$

where p_i is an individual teacher's probability of promotion, e_i^* is the teacher's optimal effort, and s_i is skill. We estimate the following equation:

$$pr_{irc,t+1} = \varphi_0 + \varphi_1 ev_{irct} + \varphi_2 p^*_{rc} + \varphi_3 n_{r\tau t} + \varphi_4 a_{ir\tau t} + \varphi_5 X_{irct} + \gamma_c + \epsilon_{irct} \tag{11}$$

where $pr_{irc,t+1}$ is a dummy variable for whether teacher i is promoted at time $t+1$ at rank r in county c, ev_{irct} is the effort (proxied by evaluation scores) of the teacher at time t, and the other variables are as defined earlier.

One difficulty in estimating the above equation is simultaneity bias caused by the fact that promotion probability is likely to induce effort in addition to being the result of greater effort. The previous regressions on the determinants of effort suggest a solution to this problem– using wage increases associated with promotion to a given rank in a given county (ΔEV) as an instrument for effort. We believe this instrumental variable to be both informative and valid. We can test the former by verifying whether the promise of a higher wage motivates teachers to work harder from the first effort regressions (equation (8)), which serves as a first stage regression. The latter is based on the assumption that wage increases for different ranks within the same county are exogenous and are not influenced by the effort of teachers or by unobservables correlated with teacher effort. China's education system is highly decentralised. It is the job of the local school board to implement the national curriculum, and county governments have substantial authority in setting teachers' wage rates. The wage rates reflect availability of fiscal revenue (Ding and Lehrer, 2007b), and can also reflect personal preferences of current leaders. Furthermore, wages affect promotions only through effort, because a teacher can access an increased wage only by working hard to get a promotion. A separate regression including various controls as well as the measure of teacher effort (annual performance evaluation scores) reveals that neither of these systematically influences wages (see Tables 3 and 4). In addition, there is no teacher's union in China, so concerns that may be present in other countries, such as high ranked teachers being able to negotiate higher wage gaps for themselves and being able to influence their likelihood of promotion, are not a concern in this context. As such, wage rates are taken to be exogenous to teachers.

From the second stage, we can test Prediction 1. The coefficient φ_1 will tell us whether teacher effort leads to promotions, and thus, whether we can expect the promotion system to motivate teacher effort.

6 Results

This section presents the empirical estimation results. We first present results for the determinants of teacher effort (equations (8) and (9)); we then present results for the determinants of promotion (equation (11)).

6.1 Effort Incentives

Table 5 presents results from estimating equation (8) for teachers whose rank levels are primary 1, middle 2, or middle 1. The outcome variable is the evaluation score of teacher i in year t. County fixed effects are controlled for, and standard errors are clustered at the county level. We find that average levels of effort are higher with higher wage gaps. The coefficient on the change in log wages from one rank to another is positive and significant. This confirms Prediction 2; that effort is higher when the prize from promotion is higher. On average, a 10% increase in the 'prize' increases the evaluation score by 0.034.

We also find evidence for Prediction 3; that average effort incentives are higher the closer is $p*$ to 1/2. Since the omitted category is a dummy variable for promotion rates in the middle quintile, we expect effort incentives to be highest here, and to be lower for both lower and higher promotion rates. The coefficients on the dummies for other promotion rate quintiles are negative as expected, with the exception of quintile 2, but the coefficients are imprecisely estimated. The coefficients on ability and the number of teachers competing are also imprecisely estimated.

The results from estimating equation (9) for effort in primary 1, middle 2 and middle 1 rank levels are reported in Table 6. The number of competitors matters. Table 6 shows that the coefficients on the number of teachers interacted with ability in the top and bottom 10% of the ability distribution, where we expect ability and the promotion rate to differ substantially, are negative, and the coefficient on the number of teachers interacted with ability in the top decile is significant. This is evidence of Prediction 5 of the model: that more competitors increases effort incentives for those with skill percentile close to one minus the promotion rate, and decreases incentives for those further away.

Figure 6 plots the coefficients on the time dummy variables from Table 6. The pattern of the dummy variables corresponds to what the model predicted: an inverted-U shape for primary 1, middle 2, and middle 1 levels. The omitted category is $t \leq X - 5$, where we expect no effort incentives. As predicted by the model (Prediction 6), from $t = X - 4$ to $t = X$ effort is increasing, apart from the first year ($t = X - 4$). As teachers get closer to eligibility for promotion, they work harder. The evaluation scores when $t = X - 2$, $t = X - 1$, and $t = X$ are significantly different than the time period $t \leq X - 5$, at the 10%, 1%, and 5% levels, respectively. The average evaluation score five years or more before eligibility ($t \leq X - 5$) is 2.5 (between a 'pass' and 'good' score). In $t = X - 2$ and $t = X - 1$, the evaluation score is more likely to be a 'good' or 'excellent' score.

Figure 6 also shows that many years after promotion eligibility, effort is decreasing. In fact, thirteen years following promotion eligibility, evaluation scores are on average almost two points lower (which is very large for a four point scale), and this effect is significant at the 1% level. This is evidence of Prediction 7; as time passes there are fewer periods in which one could enjoy the additional income from promotion, and teachers may be revising downwards the estimate of their skill percentile. As a result of these two factors, effort incentives are lower.[22] [23] [24]

We may be concerned that the increasing trend in effort before promotion eligibility may be teachers learning over time as they are in the first few years of teaching at a specific rank level. However, in the pre-eligibility period, the average experience of primary 1, middle 2 and middle 1 teachers is 14, 9 and 14 years, respectively. These are not new teachers and they are performing the same job as in previous rank levels. Other research finds that teacher learning

[22]We check that the results for the ability dummy variables are not sensitive to small changes in the cutoffs. We re-estimate the effort regressions using the top and bottom 15% of the ability distribution instead of the top and bottom 10%. The results are contained in Table 7, Column (1). The pattern of coefficeints on the time dummy variables is preserved, and the interpretation of the main results is unaffected.

[23]We also check that the results are robust to the inclusion of the number of teachers variable, since this is a variable that has been adjusted. Equation (9) is run again omitting schools with low sampling rates (below 25% and below 35%) but again with ability in the top and bottom 10% of the ability distribution. Here as well, the magnitude of the time before and after promotion eligibility dummy coefficients changes slightly but the signs remain unchanged (see Table 7 columns (2) and (3)). The coefficients on the number of teachers interacted with the top and bottom 10% of the ability distribution are still negative. Most importantly however, the pattern of the time dummy variables remains unchanged, and the interpretation of the effect of the number of teachers also remains consistent with the previous results.

[24]The regressions were also conducted separately for male and female teachers (results not reported here). No differences were found for these groups. We also look at whether Communist party membership makes a difference. Controlling for party membership does not change the results, and separating the restults by party and non-party members shows an inverted-U pattern for both, so party membership is not driving the results.

from experience comes in the initial few years of teaching (Hanushek and Rivkin, 2006).

Another challenge may be that principals manipulate evaluation scores of teachers who are nearing promotions, awarding teachers who are becoming eligible with high evaluation scores so that they can be promoted, and so the scores do not reflect true effort. Any such manipulation of evaluation scores would also produce an upward trend before eligibility, and would mean that there is systematic measurement error that is correlated with the time to eligibility for a promotion. In addition, one may question the use of evaluation scores as an accurate measure of effort, as they also reflect ability.

In order to investigate these points further, two tests of the data are performed. First, we run a Logit regression with fixed effects to examine how the probability of obtaining an 'excellent' or 'good' evaluation score[25] in 2006 is affected by the log of total time spent on teaching as well as a host of teacher controls (education dummies, age, female, where teacher is from, marital status, number of teachers, ability dummies, number of teachers interacted with ability dummies, proportion promoted quintiles, time before and after promotion eligibility dummies, and rank level dummies). The 2007 GSCF survey asks each teacher "in a typical week, how many hours do you spend on...". The information collected is for: grading homework, preparing lesson plans, participating in teaching and research activities, coaching of students outside class, organising extracurricular activities, home visits, and disciplining students. We sum up the total number of hours spent on these activities in a given week, and take the log.

This regression using the time use variable can be carried out only for 2006, since this was the only year for which time-use data was collected. The coefficient estimates are contained in Table 8. Columns (1) and (3) include only the log of total time use and columns (2) and (4) add the controls. Columns (1) and (2) include county fixed effects and columns (3) and (4) include district fixed effects. Total time use is positively and significantly related to the evaluation scores. This relationship provides strong evidence that variation in evaluation scores is not simply driven by manipulation but does reflect differences in actual effort exerted.

Second, we perform an analogous test as the one above, but with student test scores instead of time use. The results are contained in table 9. Column (1) includes county fixed effects and column (2) includes district fixed effects. The outcome is again a dummy variable for whether

[25]There is no significant difference in time use between the categories of 'excellent' and 'good' evaluation scores.

the teacher received an 'excellent' or 'good' evaluation score in 2006. The student test scores are captured by a variable defined as the average of math and language test scores for all students observed for the teacher in 2006. A control for the number of students for whom test scores are observed for each teacher is also included, as are the same controls as those listed above in the regression for time use. Student test scores are also positively and significantly related to the probability that a teacher receives an 'excellent' or 'good' evaluation score. Student test scores and time spent with students are only two aspects of teacher evaluations.

Other papers also provide evidence against the evaluation scores being manipulated. If principals were awarding high evaluation scores to those nearing promotion eligibility, all teachers would get promoted eventually and the ranks would have no meaning. There would be little or no correlation between ranks and performance. However, Park and Hannum (2001) showed that rank is a significant correlate in a regression of student test scores on various teacher characteristics using data from Gansu. Ding and Lehrer (2007a) find that teachers' ranks explain a much greater share of variation in student test scores than education or years of experience, also using Chinese data. The fact that in each rank some teachers are never promoted is also evidence that promotions are not automatic, and that evaluations are not manipulated to ensure that everyone is promoted. Finally, because evaluation scores are announced at teacher meetings and because principals are also evaluated and promoted based on the performance of their school, this should reduce the incentive to inflate evaluation scores.

Another possible type of manipulation is principals assigning teachers to easier or better classes or grades within the school so that their student's test scores are higher. However, in Gansu, in primary school there is generally only one class per grade (Park and Hannum, 2001), making such manipulation unlikely. This type of manipulation is also inconsistent with the strong relationship between time spent on teaching and test scores.

In order to test Prediction 8, we run a regression of evaluation scores on time dummy variables for primary and middle high rank levels. Primary and middle high levels are the highest rank levels a teacher can achieve in primary and middle school; there are no more promotions after this. They are difficult ranks to attain, and becoming a '*gaoji*' (high level) teacher carries a great amount of prestige. Table 10 contains results of the effort regression for primary and middle high ranks, and Figure 7 plots the time dummy variable coefficients for the number of years into the rank. The coefficients are steadily decreasing. The longer a teacher stays in these ranks, the less hard she works. This is consistent with lack of incentives in this rank, although the response is more gradual than theory might predict, perhaps reflecting gradual adjustment by teachers or the fact that these teachers are nearing retirement age and slowing down. By the time a teacher has been at primary or middle high level for 14 years, the evaluation scores are more than one point lower, and this effect is significant at the 1% level. At the primary and middle high levels, there are no predictions on the effect of ability or the number of competitors, and as in Table 10, these coefficients all are insignificant.

One also might be worried about manipuation if principals decide that excellent and good evaluations should not be 'wasted' on high rank teachers. However, similar to the results for teachers with lower ranks, teachers with primary high and middle high rank levels exhibited a pattern of time use on teaching that mirrored the evaluation scores. In this case, time spent on teaching declined with more years at the highest rank levels. We regress the time spent teaching against the number of years into primary and middle high rank with the same specification as for evaluation scores, but only for the year 2006 (since that is the year for which we have this time use data). We plot the coefficients on the number of years into the high rank level in Figure 8.[26]

6.2 Promotions

The results from estimating equation (11) for primary 1, middle 2 and middle 1 rank levels using an IV-Probit regression with an endogenous variable are contained in Table 11. This is the second stage of the two-step estimation where the first stage is contained in Table 5. The marginal effect for the evaluation score (instrumented with change in log wages) is 0.289, and this effect is significant at the 1% level. An increase in the evaluation score of one unit (for example, from 'pass' to 'good' or from 'good' to 'excellent') increases the probability of promotion by almost 30%. This result indicates that effort is being exerted by teachers in order to earn promotions, confirming Prediction 1.

There is also evidence that wages are an informative instrument. The F-statistic of 10.66

[26]This result is not driven by high rank teachers spending a disproportionate amount of time on tutoring students outside of school. Less than 10% of total time is spent on tutoring, and high rank teachers spend approximately the same proportion of time tutoring as do teachers at lower rank levels.

for the significance of the instrument reveals a strong first stage relationship, so we do not have a weak instrument problem (Staiger and Stock, 1997).

Hypothesis tests on the correlation between the error terms in the first and second stage error terms (equivalent to testing whether the endogenous variable is indeed endogenous) reveal correlations that are statistically significant. These are reported in Table 11 in the row labelled 'Rho'. The hypothesis is that the correlation is equal to zero, and we reject this at the 5% level. The sign of rho is negative, which is consistent with the difference in the evaluation score coefficients between naive Probit models and in the Probit models with IV (the coefficient is larger with IV Probit).[27]

We conduct three falsification tests to further investigate the validity of the instrument. These are reported in Table 12. First, we may be worried that wages have a direct effect on the probability of promotion. This could be the case if, for example, higher wage gaps are associated with very high or very low promotion rates and so by construction a teacher may have a higher or lower probability of promotion with a high wage gap. We can test this by looking at teacher interns at primary and middle schools. Interns are nearly automatically promoted to the lowest rank level after one year. As such, wages should not affect their effort or their promotions. In column (1), when we include the wage gap for interns directly in the promotion regression, the wage gap has no effect. The evaluation scores also do not affect the probability of promotion, as expected. We also reduce endogeneity concerns by controlling directly for historic promotion rates for each rank level in each county.

The second falsification test assigns each teacher the wage gap at the rank below, instead of the wage gap he or she currently faces. A valid instrument would mean that a wage gap that is no longer relevant should have no effect on effort. Column (2) of Table 12 reports the result. The coefficient is small and insignificant; the 'false' wage gap does not predict effort.[28]

Thirdly, we may be worried that there are unobserved factors at the rank level within a county that determine the differences in both wage gaps and evaluation scores at different levels. This could happen if, for example, the wage gap from the lowest level to the next is low and teachers in that rank tend to receive low evaluation scores. If this conjecture is true and

[27]This same pattern also holds in comparing the OLS and IV results, results not reported.

[28]We do not include the wage gap for the rank above as another falsification test because future wage increases could also motivate current teacher effort.

these patterns are similar across counties, then randomly assigning wage gaps at the same level from a different county, should also have an effect on evaluation scores. Column (3) of Table 12 shows that there is no such effect; the coefficient is small and not significantly different from zero.

7 Conclusion

In this study, we have examined the incentive properties of China's promotion system for public employees, introducing China's complex system of annual evaluations and promotions of primary and middle school teachers. As far as we are aware, this is the first paper to provide empirical evidence linking promotion incentives to work effort in the public sector, extending a limited body of previous research focused on the private sector. The Chinese system's reliance on promotion incentives also stands as an alternative model to pay for performance schemes in incentivizing teacher performance, although to date it has received little attention.

Theoretical predictions from a tournament model of promotions are largely verified by the empirical findings. Promotions are associated with significant salary increases, and teachers exert more effort, measured by annual evaluation scores, when wage increases are greater. Average effort is greater when the promotion probability is closer to 50%, and increasing the number of competitors leads to relatively less effort among those at the extremes of the skill distribution. Teacher effort increases over time as teachers get closer to the period of eligibility for promotion, and decreases if they are repeatedly passed over for promotion. Teachers who reach the highest rank exhibit a steady decline in effort. Higher annual evaluation scores significantly increase the probability of being promoted. The evaluation scores are strongly correlated with student test scores and teacher time spent on teaching, which is inconsistent with evaluation scores simply being manipulated to help teachers get promoted or for other political reasons.

The findings in this paper suggest that teachers respond to the incentives embedded in the design of the promotion system. One difference between China's promotion scheme and typical pay-for-performance schemes is that evaluations are based on a diversity of criteria and information sources (including test scores, peer and student assessments, classroom observation, evidence of innovation in educational practice or research, etc.) which are less likely than pay-for-performance schemes to lead teachers to focus only test performance at the expense of learning objectives. Unlike pay-for-performance schemes that reward recent performance, the promotion system provides dynamic incentives that may better enable teachers to set career goals and work hard to realize their career ambitions, and encourage them to make longer-term investments to becoming better teachers.

At the same time, the promotion system fails to provide strong effort incentives for some types of teachers. Those at the extreme ends of the skill distribution may perceive that their effort is unlikely to substantially influence their chances of receiving top evaluation scores or being promoted, which is especially concerning for students of the lowest ability teachers. Teachers may exert less effort if they are far from being eligible for promotion or if they are repeatedly passed over for promotion. Teachers who reach the highest rank levels may lack incentives to continue to exert high levels of effort. For such teachers, there may be a role for more immediate incentives (e.g., bonus pay) or other reforms of the existing promotion system (e.g., creating additional rank levels, shortening the time to promotion eligibility, etc.). Given the diversity of teacher ability and effort incentives, it could be optimal to combine immediate performance incentives with promotion incentives as argued by Kwon (2006).

Although the specific promotion system that we study only directly affects the performance of primary and middle school teachers in China, it shares many elements with the promotion systems governing all public servants in China. For this reason, it can help shed light on the puzzle of how China has grown so fast despite many institutional limitations suggesting poor governance (e.g., corruption, lack of property rights, poor credit access for private firms). One area in which China may be doing much better than other countries is providing strong effort incentives to government officials at all levels.

References

Acemoglu, D., J. A. Robinson, and D. Woren (2012). *Why nations fail: the origins of power, prosperity, and poverty*, Volume 4. SciELO Chile.

Ashraf, N., O. Bandiera, and S. S. Lee (2014a). Awards unbundled: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization 100*, 44–63.

Ashraf, N., O. Bandiera, and S. S. Lee (2014b). Do-gooders and go-getters: career incentives, selection, and performance in public service delivery. Technical report, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Baker, G., M. J. Gibbs, and B. Holmstrom (1994a). The internal economics of the firm: evidence from personnel data. *Quarterly Journal of Economics 109*(4), 881–919.

Baker, G., M. J. Gibbs, and B. Holmstrom (1994b). The wage policy of a firm. *Quarterly Journal of Economics 109*(4), 921–955.

Barlevy, G. and D. Neal (2012). Pay for Percentile. *American Economic Review 102*(5), 1805–1831.

Bruns, B., D. Filmer, and H. A. Patrinos (2011). *Making schools work: New evidence on accountability reforms*. World Bank Publications.

Campbell, D. (2008). Nonfinancial performance measures and promotion-based incentives. *Journal of Accounting Research 46*(2), 297–332.

Chan, K. W., V. Henderson, and K. Y. Tsui (2008). Spatial Dimensions of Chinese Economic Development. In T. Rawski and L. Brandt (Eds.), *China's Great Transformation: Origins, Mechanisms, and Consequences of the Post-Reform Economic Boom*, pp. 776–828. Cambridge University Press.

Dal Bó, E., F. Finan, and M. A. Rossi (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *The Quarterly Journal of Economics 128*(3), 1169–1218.

Ding, W. and S. F. Lehrer (2007a). Do peers affect student achievement in China's secondary schools? *The Review of Economics and Statistics 89*(2), 300–312.

Ding, W. and S. F. Lehrer (2007b). The optimal policy to reward the value added by educators: Theory and evidence from China. In E. Hannum and A. Park (Eds.), *Education and Reform in China*. Routledge.

Dixit, A. (1997). Power of incentives in private versus public organizations. *American Economic Review 87*(2), 378–382.

Dixit, A. (2002). Incentives and organizations in the public sector: An interpretative review. *Journal of Human Resources*, 696–727.

Doeringer, P. B. and M. J. Piore (1971). *Internal labor markets and manpower analysis.* ME Sharpe.

Ederer, F. (2010). Feedback and motivation in dynamic tournaments. *Journal of Economics & Management Strategy 19*(3), 733–769.

Enikolopov, R. (2011). Are bureaucrats really paid like bureaucrats? Technical report.

Fairburn, J. A. and J. M. Malcomson (2001). Performance, promotion, and the Peter Principle. *The Review of Economic Studies 68*(1), 45–66.

Ferraz, C. and F. Finan (2009). Motivating politicians: The impacts of monetary incentives on quality and performance. Technical report, National Bureau of Economic Research.

Figlio, D. N. and L. S. Getzler (2006). Accountability, ability and disability: Gaming the system? *Advances in Applied Microeconomics 14*, 35–49.

Figlio, D. N. and J. Winicki (2005). Food for thought: the effects of school accountability plans on school nutrition. *Journal of Public Economics 89*(2), 381–394.

Gibbs, M. J. (1989). *Promotions, compensation, and firm organization.* Ph. D. thesis, University of Chicago, Department of Economics.

Gibbs, M. J. (1995). Incentive compensation in a corporate hierarchy. *Journal of Accounting and Economics 19*(2), 247–277.

Glewwe, P. (2002). Schools and skills in developing countries: Education policies and socioeconomic outcomes. *Journal of Economic Literature 40*(2), 436–482.

Glewwe, P., N. Ilias, and M. Kremer (2010). Teacher incentives. *American Economic Journal: Applied Economics 2*(3), 205–227.

Haeck, C. and F. Verboven (2012). The internal economics of a university: Evidence from personnel data. *Journal of Labor Economics 30*(3), 591–626.

Han, L. (2013). Is centralized teacher deployment more equitable? Evidence from rural China. *China Economic Review 24*, 65–76.

Han, L., M. Liu, and X. An (2014). Centralized deployment and teachers incentive: Evidence from reforms in rural China. *mimeo, HKUST*.

Hansen, S. E. (2012). Performance feedback with career concerns. *Journal of Law, Economics, and Organization*.

Hanushek, E. A. and S. G. Rivkin (2006). Teacher quality. *Handbook of the Economics of Education 2*, 1051–1078.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 24–52.

Jacob, B. A. and S. D. Levitt (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics 118*(3), 843–877.

Kwon, I. (2006). Incentives, wages, and promotions: theory and evidence. *RAND Journal of Economics 37*(1), 100–120.

Landry, P. F. (2008). *Decentralized Authoritarianism in China: the Communist Party's control of local elites in the post-Mao era.* Cambridge University Press New York.

Lavy, V. (2003). Paying for performance: The effect of individual financial incentives on teachers productivity and students scholastic outcomes. *Jerusalem, Hebrew University*.

Li, H. and L.-A. Zhou (2005). Political turnover and economic performance: the incentive role of personnel control in China. *Journal of Public Economics 89*(9), 1743–1762.

Macartney, H. (2014). The dynamic effects of educational accountability. Technical report, National Bureau of Economic Research.

Medoff, J. L. and K. G. Abraham (1980). Experience, performance, and earnings. *Quarterly Journal of Economics 95*(4), 703–736.

Medoff, J. L. and K. G. Abraham (1981). Are those paid more really more productive? The case of experience. *Journal of Human Resources*, 186–216.

Muralidharan, K. (2011). Long-term effects of teacher performance pay: Experimental evidence from India. *mimeo*.

Muralidharan, K. and V. Sundararaman (2011). Teacher performance pay: Experimental evidence from india. *Journal of Political Economy 119*(1), pp. 39–77.

Neal, D. (2011). The design of performance pay in education. In E. A. Hanushek, S. J. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 4. Elsevier.

Neal, D. and D. W. Schanzenbach (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics 92*(2), 263–283.

OECD (2010). *PISA 2009 Results: What Students Know and Can Do: Student Performance in Reading, Mathematics and Science*. OECD.

Park, A. and E. Hannum (2001). Do teachers affect learning in developing countries? Evidence from matched student-teacher data from China. In *Conference Rethinking Social Science Research on the Developing World in the 21st Century, Park City, Utah, June*, pp. 7–11.

Podgursky, M. J. and M. G. Springer (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management 26*(4), 909–950.

Propper, C. and J. Van Reenen (2010). Can pay regulation kill? panel data evidence on the effect of labor markets on hospital performance. *Journal of Political Economy 118*(2), 222–273.

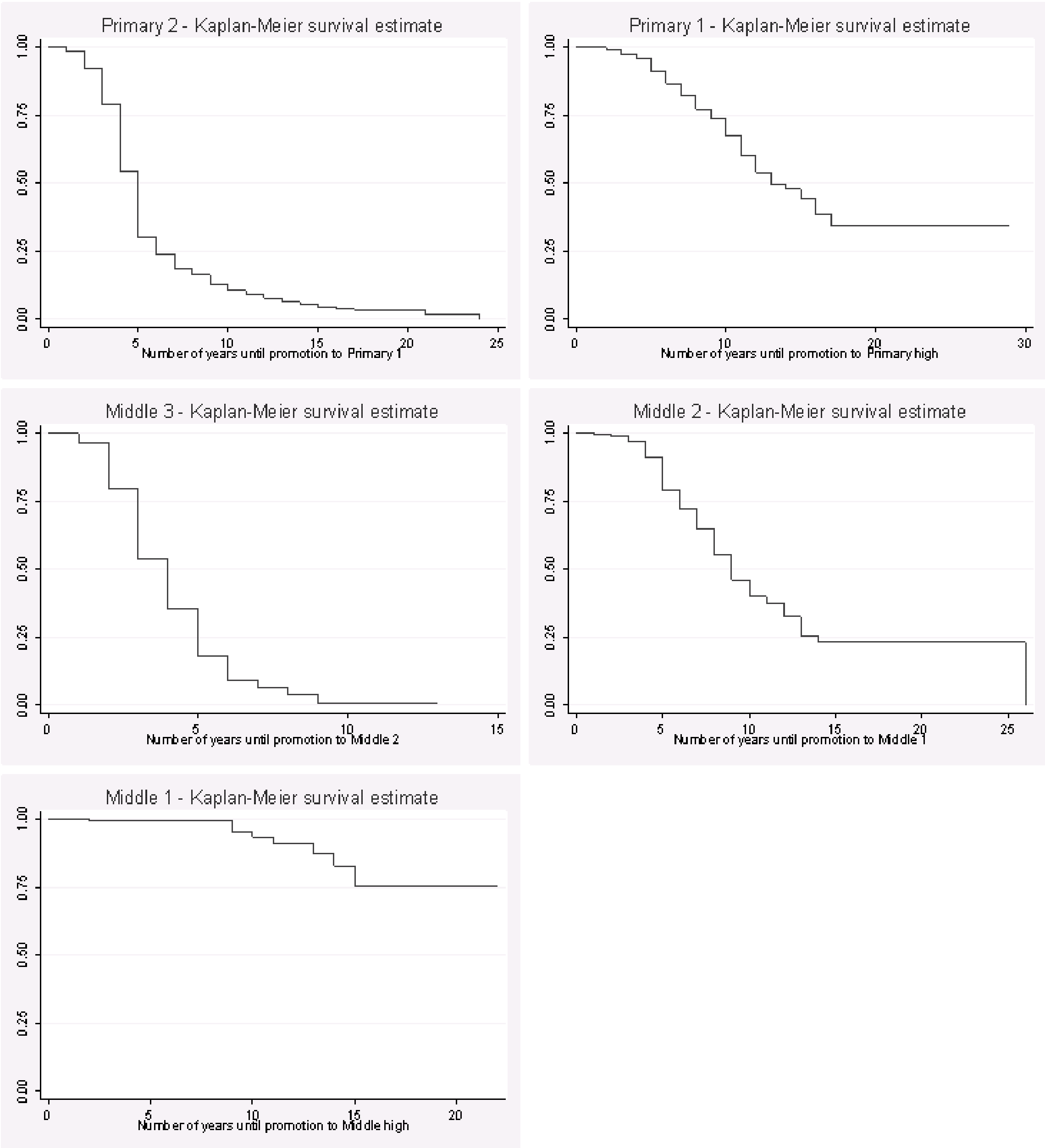
Rasul, I. and D. Rogger (2013). Management of bureaucrats and public service delivery: Evidence from the nigerian civil service. *Working Paper, University College London*.

Staiger, D. and J. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*, 557–586.

World Bank (2004). World Development Report: Making Services Work for Poor People. Technical report, World Bank.

A Appendix: Figures

Figure 1: Kaplan-Meier Survival Functions



Source: Gansu 2007 Data.

Figure 2: Effort Incentives: $p^* = 1/2$, $p^* < 1/2$, $p^* > 1/2$

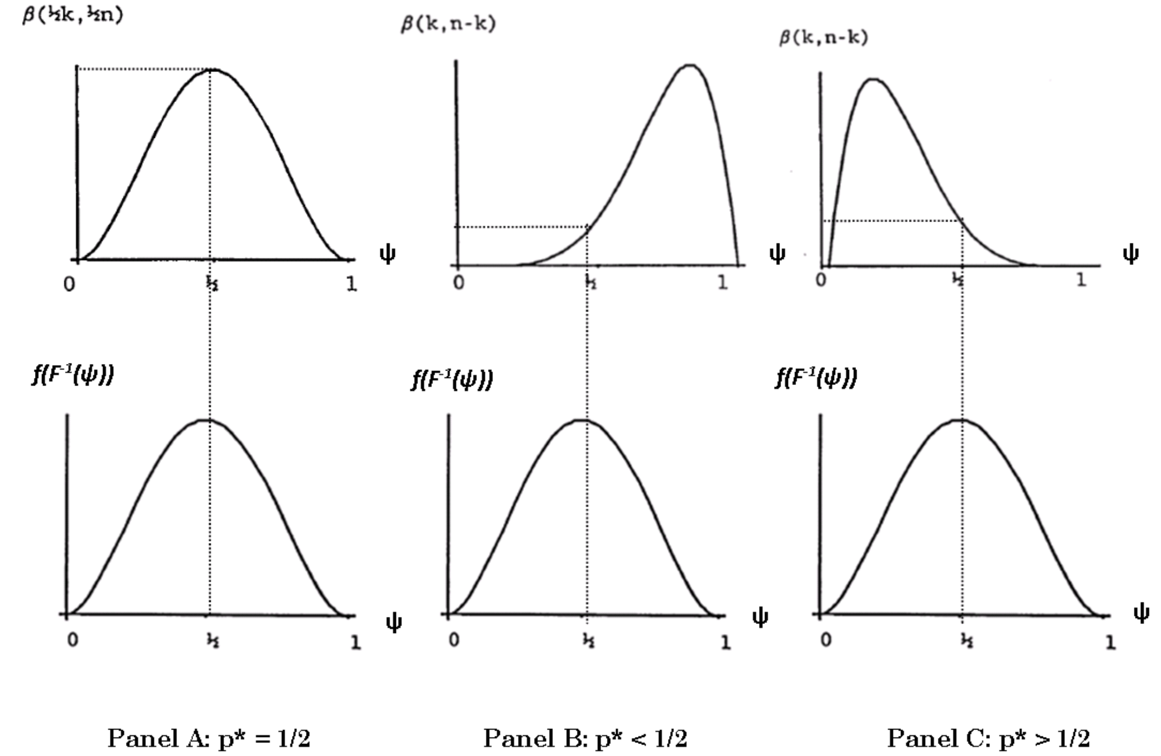


Figure 3: The marginal-skilled teacher

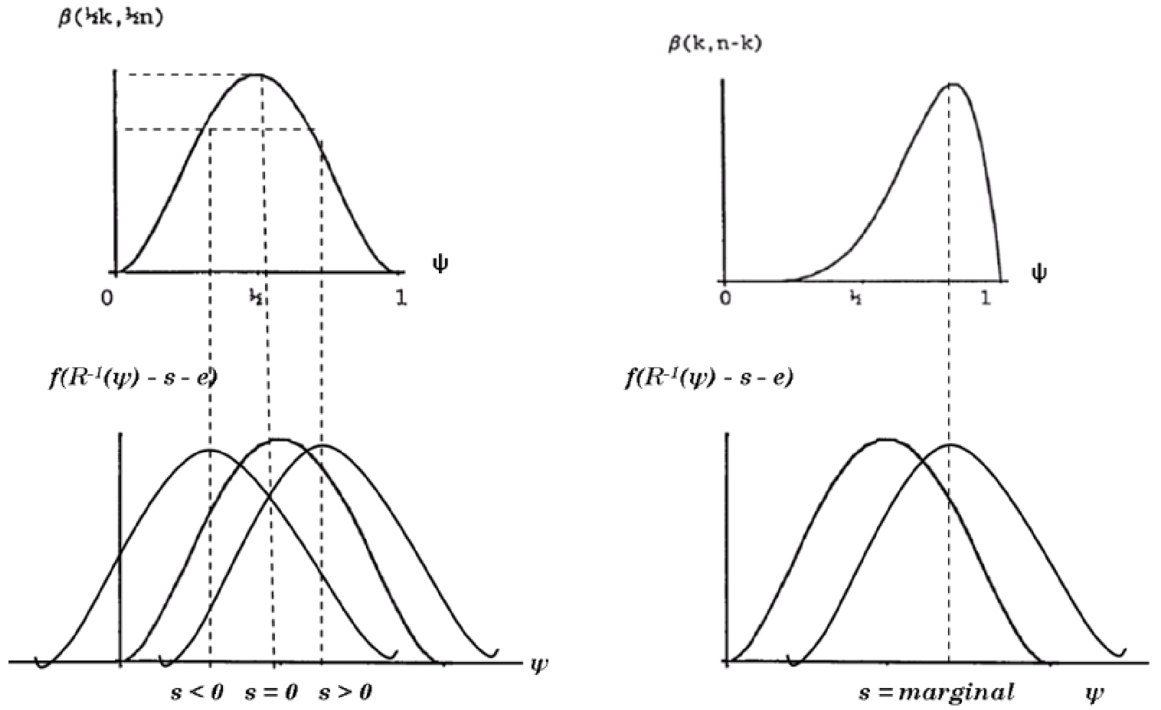


Figure 4: Heterogeneous competitors and number competing

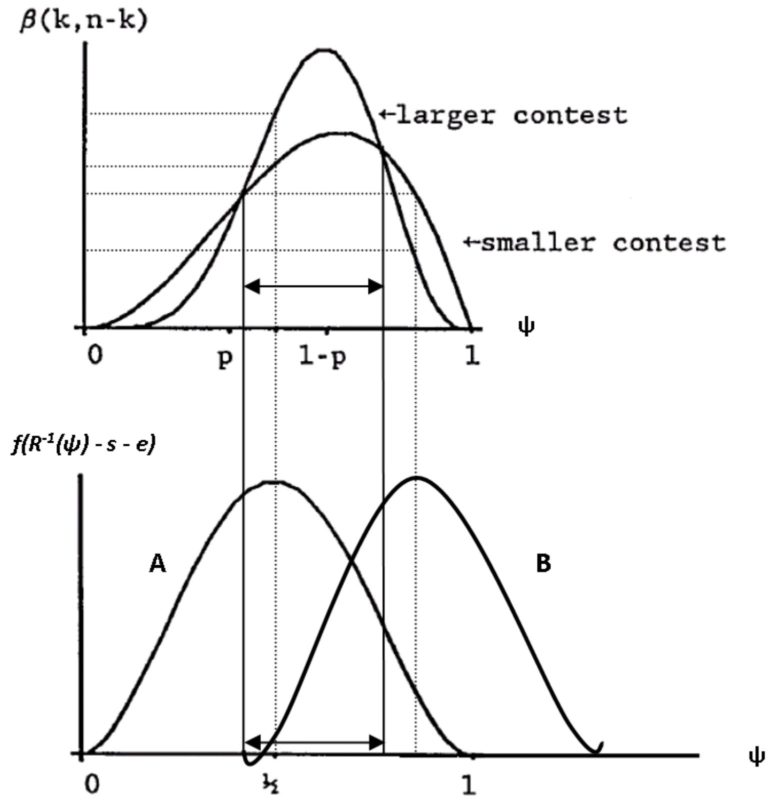


Figure 5: Effort incentives over time

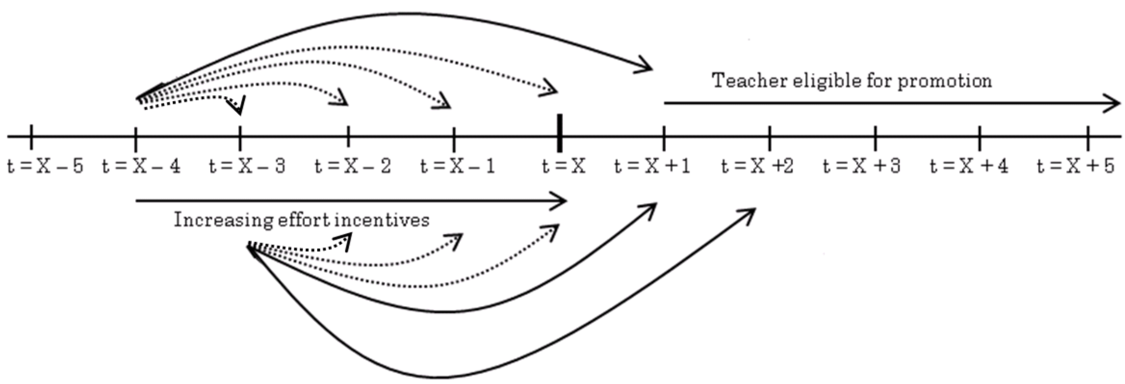
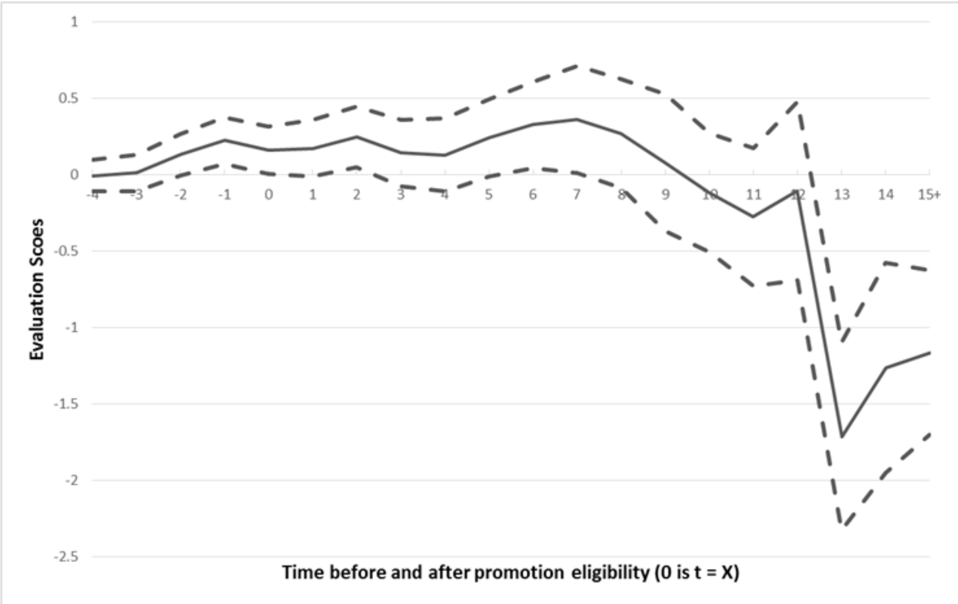
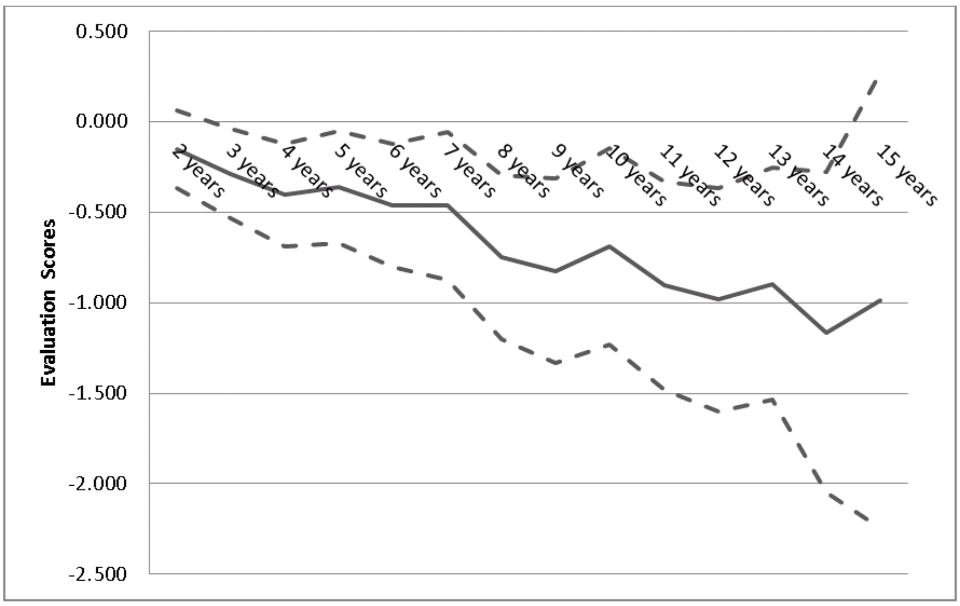


Figure 6: Primary 1, middle 2 and middle 1, pre- and post-eligibility time dummies



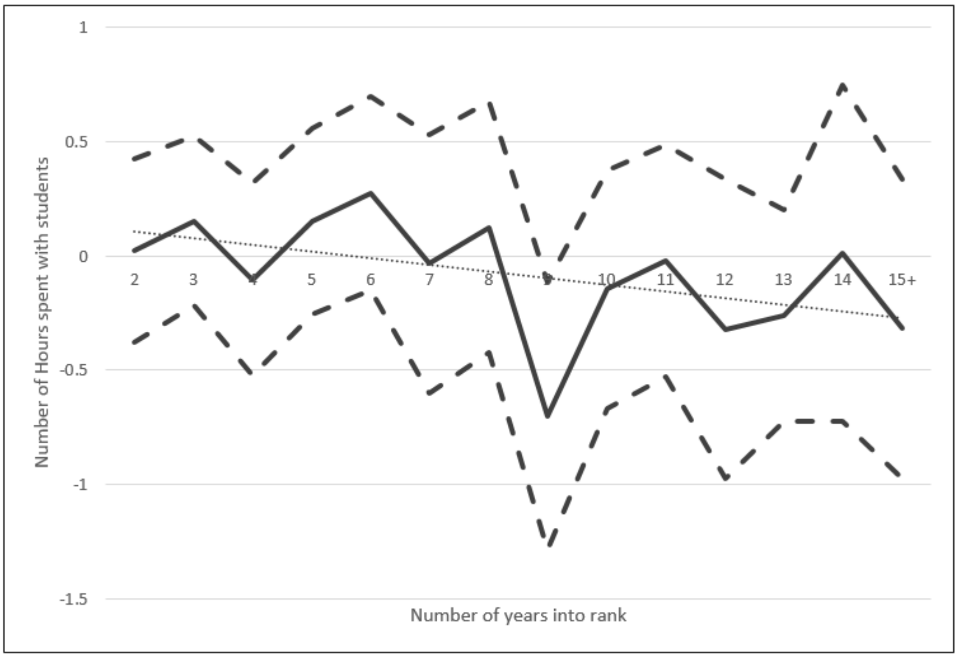
Notes: Dashed lines are 95% confidence intervals.

Figure 7: Primary and middle high time dummies



Note: Dashed lines are 95% confidence intervals.

Figure 8: Primary and middle high hours spent with students



B Appendix: Tables

Table 1: Characteristics of teachers in the sample, Round 3 (2007)

	Intern	Primary 2	Primary 1	Primary high	Middle 3	Middle 2	Middle 1	Middle high	All teachers
Number of teachers									
Total	135	141	544	348	127	516	277	12	2,100
Proportion Female	0.53	0.57	0.45	0.25	0.50	0.37	0.17	0.17	0.38
Basic characteristics									
Average Age	25.38	27.45	35.65	46.99	25.87	31.28	39.67	45.75	35.25
	(6.45)	(8.22)	(8.92)	(6.59)	(3.68)	(5.33)	(6.11)	(5.28)	(9.67)
Average years teaching	5.88	6.31	15.27	26.64	3.33	9.04	18.41	26.25	14.62
	(9.55)	(7.90)	(8.91)	(6.02)	(3.62)	(5.13)	(6.58)	(4.56)	(9.92)
Education - primary	1%	1%	0%	0%	0%	0%	0%	0%	0.2%
Education - middle school	5%	2%	3%	3%	1%	2%	1%	0%	2.5%
Education - high school	23%	7%	22%	48%	2%	2%	10%	0%	17.6%
Education - vocational middle	20%	55%	63%	45%	13%	32%	47%	54%	42.3%
Education - vocational tech	1%	1%	0%	0%	0%	0%	1%	0%	0.3%
Education - vocational college	45%	34%	12%	4%	83%	57%	38%	46%	34.2%
Education - four year degree	6%	1%	0%	0%	2%	7%	3%	0%	3%
Time use									
Number of classes taught	15.63	19.12	18.03	17.73	14.34	13.66	13.11	10.17	15.91
	(5.89)	(5.67)	(5.87)	(6.65)	(5.11)	(4.03)	(4.27)	(4.69)	(5.80)
Hours per week spent grading	10.62	11.99	9.63	9.26	10.09	9.12	9.12	7.50	9.61
	(5.33)	(6.65)	(4.87)	(5.48)	(6.64)	(4.72)	(4.56)	(4.98)	(5.24)
Hours per week spent on lesson plans	9.83	9.91	9.31	9.15	9.83	9.18	9.69	8.25	9.40
	(6.43)	(5.06)	(5.07)	(5.70)	(5.84)	(4.92)	(4.93)	(5.63)	(5.27)
Hours per week spent on teaching	3.12	3.38	2.89	2.74	2.73	2.90	2.81	2.00	2.89
	(2.87)	(2.68)	(2.13)	(2.22)	(2.31)	(2.95)	(1.93)	(1.13)	(2.45)
Hours per week spent on tutoring	5.33	5.59	4.75	4.40	5.87	4.97	4.70	2.83	4.89
	(3.43)	(3.79)	(3.28)	(3.86)	(3.98)	(3.74)	(3.60)	(2.08)	(3.64)
Hours per week spent on extra-curriculars	3.31	3.26	2.98	2.83	2.65	2.59	2.51	1.58	2.81
	(4.56)	(2.50)	(2.21)	(2.73)	(2.30)	(2.27)	(1.89)	(1.83)	(2.53)
Hours per week spent on home visits	2.35	2.35	1.89	1.76	2.53	1.96	1.80	0.63	1.97
	(3.65)	(2.23)	(2.48)	(1.63)	(2.83)	(2.34)	(2.15)	(0.64)	(2.39)
Hours per week spent on discipline	2.94	2.73	2.51	2.27	3.48	3.23	3.29	2.08	2.85
	(2.90)	(1.78)	(2.56)	(2.02)	(3.94)	(2.84)	(3.25)	(2.02)	(2.76)
Total hours per week	52.99	58.05	51.78	50.06	51.19	47.55	46.81	35.04	50.17
	(20.73)	(18.79)	(15.89)	(20.37)	(19.78)	(16.15)	(15.46)	(16.06)	(17.75)
Published a paper	12.1%	25.2%	46.5%	35.9%	17.3%	45.9%	57.3%	100.0%	38.3%
Wages									
Monthly wage (CNY)	972.27	1,034.05	1,296.66	1,517.26	1,038.24	1,274.63	1,535.41	1,880.75	1,308.45
	(256.97)	(204.76)	(218.35)	(262.84)	(208.42)	(222.94)	(236.50)	(269.30)	(293.84)
Log of monthly wage (CNY)	6.84	6.92	7.16	7.31	6.92	7.13	7.32	7.53	7.15
	(0.30)	(0.21)	(0.17)	(0.17)	(0.23)	(0.19)	(0.16)	(0.14)	(0.24)
Variance of log wages	0.092	0.045	0.029	0.030	0.051	0.035	0.024	0.019	0.058

Note: Interns include Primary and Middle school interns. Standard deviations in parentheses.

Table 2: Evaluation scores of teachers by rank (2003-2006)

	Interns		Primary 2		Primary 1		Primary high		Middle 3		Middle 2		Middle 1		Middle high	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Fail	-	-	3	0.4	4	0.2	1	0.1	-	-	5	0.3	1	0.1	-	-
Pass	236	43.62	396	48.4	973	49.9	526	50.7	473	56.1	906	54.9	802	51.1	16	55.2
Good	215	39.74	291	35.5	675	34.6	378	36.5	253	30.0	487	29.5	502	32.0	10	34.5
Excellent	90	16.64	129	15.8	297	15.2	132	12.7	117	13.9	251	15.2	263	16.8	3	10.3
Total	541	100	819	100	1,949	100	1,037	100	843	100	1,649	100	1,568	100	29	100

Note: Interns include Primary and Middle school interns.

Table 3: Wage Premia - Primary School

Outcome: log of monthly wage, 2007	(1)	(2)	(3)
Primary 1	0.204***	0.084***	0.075***
	(0.030)	(0.023)	(0.024)
Primary high	0.361***	0.150***	0.132***
	(0.034)	(0.030)	(0.033)
Experience		0.021***	0.022***
		(0.004)	(0.004)
Experience squared		-3.23E-04***	-3.03E-04***
		(9.79E-05)	(9.37E-05)
Education - middle school			0.038
			(0.025)
Education - high school			0.048***
			(0.011)
Education - vocational middle school			0.099***
			(0.015)
Education - vocational college			0.077
			(0.050)
Education - four year degree			0.089***
			(0.021)
Education - graduate school			0.190**
			(0.074)
Year end assessment result 2006			0.001
			(0.005)
Observations	1,029	1,029	1,026
R^2	0.400	0.543	0.554

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include county fixed effects and clustered standard errors at the county level. Performance evaluation score is measured on a scale of 1-4 (fail, pass, good, excellent).

Table 4: Wage Premia - Middle School

Outcome: log of monthly wage, 2007	(1)	(2)	(3)
Middle 2	0.134***	0.047*	0.039*
	(0.030)	(0.024)	(0.021)
Middle 1	0.318***	0.127***	0.117***
	(0.035)	(0.025)	(0.025)
Middle high	0.535***	0.301***	0.280***
	(0.080)	(0.063)	(0.058)
Experience		0.021***	0.022***
		(0.004)	(0.004)
Experience squared		-3.30E-04***	-3.43E-04***
		(9.64E-05)	(1.11E-04)
Education - high school			-0.036
			(0.040)
Education - vocational middle school			-0.017
			(0.018)
Education - vocational college			0.089
			(0.112)
Education - four year degree			3.925E-04
			(0.022)
Education - graduate school			0.012
			(0.028)
Year end assessment result 2006			-0.003
			(0.005)
Observations	930	930	926
R^2	0.390	0.493	0.494

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include county fixed effects and clustered standard errors at the county level. Performance evaluation score is measured on a scale of 1-4 (fail, pass, good, excellent).

Table 5: Parameters of the promotion system and effort incentives

Outcome: evaluation score of teacher i in year t		
	(1)	
	Primary 2, Middle 2, Middle 1	
	coefficient	se
Change in log wages	0.877***	0.269
Education - high school	-0.164**	0.080
Education - vocational college	0.004	0.039
Education - college	0.075*	0.044
Experience (years)	0.010***	0.002
Female	-0.036	0.028
Proportion promoted - quintile 1	-0.013	0.044
Proportion promoted - quintile 2	0.022	0.047
Proportion promoted - quintile 4	-0.108**	0.046
Proportion promoted - quintile 5	-0.059	0.047
Ability in bottom 10%	-0.050	0.039
Ability in top 10%	0.086**	0.042
Number of teachers in same rank and district (log)	-2.212E-04	1.947E-04
Number of teachers (log) * ability in bottom 10%	-2.818E-04	3.850E-04
Number of teachers (log) * ability in top 10%	4.922E-04	4.861E-04
Constant	2.325***	0.080
Number of observations	3,441	

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regression includes county fixed effects and clustered standard errors at the county level. Proportion promoted is the proportion of teachers in a county that are promoted at that rank within a 'reasonable' number of years. We first calculate the time it took for each teacher to be promoted from one rank to another. Then we calculate the number of years for which, for each rank and education level in a county, half of the teachers were promoted within that time. Then for each rank and county we calculate the proportion of teachers promoted within that 'reasonable' time. We only include promotions before 2003. The ability index was constructed using the teachers' scores on an education test, a language test and a psychology test that were taken before the teacher began teaching. Scores are first standardized by the year in which they were taken, and then for each year the following measure is calculated: $[S - ave(S)]/ave(S)$ where S is the standardized score on the tests for a teacher, and $ave(S)$ is the average of the scores of the teachers in the same rank in the same school, to capture relative ability. The index was constructed using Principal Components Analysis (PCA). Dummies for the top and bottom 10% of this index are included. Evaluation score is measured on a scale of 1-4 (fail, pass, good, excellent).

Table 6: Primary 1, Middle 2, and Middle 1 rank effort incentives

Outcome: evaluation score of teacher i in year t	coefficient	standard error
Relative age	0.285	(0.374)
Log number of teachers in rank in district (N)	0.013	(0.082)
N * ability top 10%	-0.091*	(0.052)
N * ability bottom 10%	-0.043	(0.062)
Ability in top 10%	0.320*	(0.193)
Ability in bottom 10%	0.038	(0.200)
t = X - 4	-0.008	(0.052)
t = X - 3	0.010	(0.060)
t = X - 2	0.130*	(0.068)
t = X - 1	0.223***	(0.077)
t = X	0.159**	(0.080)
t = X + 1	0.173*	(0.093)
t = X + 2	0.246**	(0.100)
t = X + 3	0.140	(0.110)
t = X + 4	0.129	(0.121)
t = X + 5	0.242*	(0.128)
t = X + 6	0.326**	(0.144)
t = X + 7	0.359**	(0.179)
t = X + 8	0.269	(0.182)
t = X + 9	0.078	(0.228)
t = X + 10	-0.117	(0.199)
t = X + 11	-0.277	(0.230)
t = X + 12	-0.106	(0.299)
t = X + 13	-1.715***	(0.313)
t = X + 14	-1.265***	(0.350)
t $\geq X + 15$	-1.165***	(0.274)
Observations	3683	
R^2	0.022	

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regression includes teacher fixed effects and clustered standard errors at the teacher level. The ability index was constructed using the teachers' scores on an education test, a language test and a psychology test that were taken before the teacher began teaching. Scores are first standardized by the year in which they were taken, and then for each year the following measure is calculated: $[S - ave(S)]/ave(S)$ where S is the standardized score on the tests for a teacher, and $ave(S)$ is the average of the scores of the teachers in the same rank in the same school, to capture relative ability. The index was constructed using Principal Components Analysis (PCA). Dummies for the top and bottom 10% of this index are included. Evaluation score is measured on a scale of 1-4 (fail, pass, good, excellent). t denotes time, X denotes the year in which a teacher becomes eligible to apply for promotion.

Table 7: Robustness Checks

	Outcome: evaluation score of teacher i in year t					
	(1)		(2)		(3)	
	Ability in top & bottom 15%		Teacher sampling rate $\geq$ 25%		Teacher sampling rate $\geq$ 35%	
Relative age	0.300	(0.317)	0.517	(0.337)	0.426	(0.342)
Log number of teachers in rank in district (N)	0.033	(0.070)	-0.022	(0.072)	-0.040	(0.075)
Ability in top 15%	0.275*	(0.165)				
Ability in bottom 15%	0.221	(0.166)				
N * ability top 15%	-0.091*	(0.051)				
N * ability bottom 15%	-0.067	(0.051)				
Ability in top 10%			0.288*	(0.172)	0.117	(0.183)
Ability in bottom 10%			0.058	(0.189)	0.033	(0.199)
N * ability top 10%			-0.078	(0.050)	-0.034	(0.052)
N * ability bottom 10%			-0.059	(0.060)	-0.044	(0.062)
t = X - 4	-0.008	(0.048)	-0.004	(0.050)	-0.017	(0.051)
t = X - 3	0.009	(0.051)	0.032	(0.054)	0.020	(0.055)
t = X - 2	0.130**	(0.056)	0.105*	(0.059)	0.108*	(0.060)
t = X - 1	0.229***	(0.062)	0.193***	(0.065)	0.197***	(0.067)
t = X	0.165**	(0.067)	0.165**	(0.071)	0.196***	(0.073)
t = X + 1	0.178**	(0.076)	0.174**	(0.080)	0.199**	(0.081)
t = X + 2	0.249***	(0.083)	0.262***	(0.086)	0.302***	(0.088)
t = X + 3	0.144	(0.090)	0.160*	(0.094)	0.200**	(0.096)
t = X + 4	0.133	(0.101)	0.117	(0.105)	0.161	(0.107)
t = X + 5	0.243**	(0.110)	0.214*	(0.114)	0.282**	(0.116)
t = X + 6	0.318***	(0.122)	0.296**	(0.128)	0.355***	(0.130)
t = X + 7	0.354**	(0.144)	0.362**	(0.149)	0.465***	(0.152)
t = X + 8	0.258	(0.161)	0.184	(0.169)	0.295*	(0.173)
t = X + 9	0.067	(0.187)	0.040	(0.189)	0.158	(0.194)
t = X + 10	-0.142	(0.210)	-0.159	(0.212)	-0.052	(0.214)
t = X + 11	-0.276	(0.341)	-0.320	(0.339)	-0.204	(0.340)
t = X + 12	-0.110	(0.343)	-0.095	(0.362)	0.024	(0.363)
t = X + 13	-1.674**	(0.798)	-1.627*	(0.971)	-1.454	(0.969)
t = X + 14	-1.261*	(0.732)	-1.944**	(0.781)	-1.885**	(0.779)
$t \geq X + 15$	-1.185*	(0.708)	-1.113	(0.713)	-1.069	(0.712)
Observations	3683		3373		3279	
R^2	0.022		0.023		0.024	

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regression includes teacher fixed effects and clustered standard errors at the level of the teacher. The ability index was constructed using the teachers' scores on an education test, a language test and a psychology test that were taken before the teacher began teaching. Scores are first standardized by the year in which they were taken, and then for each year the following measure is calculated: $[S - ave(S)]/ave(S)$ where S is the standardized score on the tests for a teacher, and $ave(S)$ is the average of the scores of the teachers in the same rank in the same school, to caputure relative ability. The index was constructed using Principal Components Analysis (PCA). Dummies for the top and bottom 10% of this index are included. Evaluation score is measured on a scale of 1-4 (fail, pass, good, excellent). t denotes time, X denotes the year in which a teacher becomes eligible to apply for promotion.

Table 8: Probability of Receiving an 'Excellent' or 'Good' evaluation score

	Outcome: excellent or good evaluation score dummy							
	(1)		(2)		(3)		(4)	
	coefficient	se	coefficient	se	coefficient	se	coefficient	se
Log of total time use	0.322**	0.162	0.381**	0.189	0.344**	0.166	0.357*	0.195
Education - vocational college			0.024	0.201			-0.077	0.208
Education - college or higher			0.149	0.224			0.086	0.235
Female			-0.100	0.141			-0.073	0.148
Age			0.013	0.011			0.014	0.011
From same village			1.120	0.826			1.093	0.826
From same town			1.182	0.825			1.242	0.825
From same county			1.055	0.819			1.109	0.819
From same province			1.240	0.852			1.204	0.854
Married			0.261	0.381			0.319	0.383
Divorced			-1.676	1.129			-1.549	1.139
Widowed			0.402	0.846			0.519	0.867
Ability in bottom 10%			-0.146	0.162			-0.151	0.171
Ability in top 10%			0.130	0.161			0.109	0.169
County fixed effects	Yes		Yes		No		No	
District fixed effects	No		No		Yes		Yes	
Number of observations	1,376		1,084		1,363		1,064	

Notes: $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the level of the fixed effect. The ability index was constructed using the teachers' scores on an education test, a language test and a psychology test that were taken before the teacher began teaching. Scores are first standardized by the year in which they were taken, and then for each year the following measure is calculated: $[S - ave(S)]/ave(S)$ where S is the standardized score on the tests for a teacher, and $ave(S)$ is the average of the scores of the teachers in the same rank in the same school, to capture relative ability. The index was constructed using Principal Components Analysis (PCA). Dummies for the top and bottom 10% of this index are included. Log of total time use first sums up the number of hours per week spent on grading homework, preparing lesson plans, participating in teaching and research activities, coaching of students outside class, organising extracurricular activities, home visits, and disciplining students, and then takes the log.

Table 9: Evaluation scores and test scores

Outcome: excellent or good evaluation score dummy		
	(1)	(2)
Marginal effect of average test scores	0.010**	0.009**
Standard error	(0.005)	(0.004)
Number of observations	258	258
Fixed effect	County	District

Notes: $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the level of the fixed effect. Controls include: number of child test scores observed for teacher, education dummies, age, female, where teacher is from, marital status, number of teachers, ability dummies, number of teachers interacted with ability dummies, proportion promoted quintiles, time before and after promotion eligibility dummies, and rank level dummies. Student test scores are defined as the average of math and language test scores for all students observed for the teacher in 2006.

Table 10: Primary and Middle high rank effort incentives

Outcome: evaluation score of teacher i in year t		
	coefficient	standard error
Relative age	-1.218	(0.763)
Log number of teachers in rank in district (N)	-0.267	(0.309)
N * ability top 10%	0.139	(0.229)
N * ability bottom 10%	0.229	(0.342)
Ability in top 10%	-0.554	(0.798)
Ability in bottom 10%	-1.124	(1.234)
2 years in rank	-0.118	(0.107)
3 years in rank	-0.240*	(0.124)
4 years in rank	-0.350**	(0.142)
5 years in rank	-0.308**	(0.156)
6 years in rank	-0.407**	(0.171)
7 years in rank	-0.407**	(0.207)
8 years in rank	-0.687***	(0.230)
9 years in rank	-0.758***	(0.258)
10 years in rank	-0.619**	(0.275)
11 years in rank	-0.835***	(0.291)
12 years in rank	-0.912***	(0.312)
13 years in rank	-0.823**	(0.324)
14 years in rank	-1.089**	(0.450)
15 + years in rank	-0.912	(0.640)
Observations	648	
R^2	0.071	

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regression includes teacher fixed effects and clustered standard errors at the teacher level. The ability index was constructed using the teachers' scores on an education test, a language test and a psychology test that were taken before the teacher began teaching. Scores are first standardized by the year in which they were taken, and then for each year the following measure is calculated: $[S - ave(S)]/ave(S)$ where S is the standardized score on the tests for a teacher, and $ave(S)$ is the average of the scores of the teachers in the same rank in the same school, to capture relative ability. The index was constructed using Principal Components Analysis (PCA). Dummies for the top and bottom 10% of this index are included. Evaluation score is measured on a scale of 1-4 (fail, pass, good, excellent).

Table 11: Promotion probability - marginal effect

Outcome: Promotion of teacher i in year t		
	(1)	
	Primary 2, Middle 2, Middle 1	
	marginal effect	se
Evaluation score (instrumented with change in log wages)	0.289***	0.065
Number of observations	3,441	
Rho	-0.937***	
Sigma	0.724***	
F statistic	10.66	
R^2	0.0387	

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regression includes county fixed effects and clustered standard errors at the county level. Regression includes same controls as Table 5.

Table 12: IV Falsification Tests

Outcome	(1) Promotion Dummy	(2) Evaluation Score	(3) Evaluation Score
Evaluation score	-0.003		
	(0.004)		
Change in log wages (interns)	0.014		
	(0.013)		
Change in log wages (of rank below)		0.176	
		(0.202)	
Change in log wages (randomly assigned counties)			-0.286
			(0.581)
Education - high school	0.001	-0.154	-0.144
	(0.002)	(0.123)	(0.116)
Education - vocational college	3.37E-04	0.007	-0.020
	(0.003)	(0.036)	(0.041)
Education - college	-0.009	0.080*	0.038
	(0.011)	(0.042)	(0.043)
Experience (years)	-3.65E-04	0.008***	0.007**
	(3.99E-04)	(0.002)	(0.003)
Female	-0.013	-0.031	-0.030
	(0.013)	(0.033)	(0.037)
Proportion promoted - quintile 1	-0.001	0.086	0.046
	(0.003)	(0.062)	(0.069)
Proportion promoted - quintile 2	-0.002	0.091	0.091*
	(0.002)	(0.058)	(0.054)
Proportion promoted - quintile 4	0.001	-0.081	-0.101*
	(0.004)	(0.052)	(0.055)
Proportion promoted - quintile 5	(dropped)	0.029	-0.010
		(0.055)	(0.044)
Ability in bottom 10%	0.015	-0.031	-0.070
	(0.015)	(0.047)	(0.050)
Ability in top 10%	-0.016	0.073	0.067
	(0.018)	(0.046)	(0.046)
Number of teachers in same rank and district (log)	0.004	-2.50E-04	-2.22E-04
	(0.004)	(3.38E-04)	(3.52E-04)
Number of teachers (log) * ability in bottom 10%	-0.006	-3.35E-04	-1.71E-04
	(0.006)	(3.72E-04)	(3.57E-04)
Number of teachers (log) * ability in top 10%	0.007	0.001	0.001
	(0.007)	(0.001)	(0.001)
Constant	-0.002	2.324***	2.467***
	(0.007)	(0.062)	(0.137)
Number of observations	371	3,588	3,507

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regression includes county fixed effects and clustered standard errors at the county level. Proportion promoted is the proportion of teachers in a county that are promoted at that rank within a 'reasonable' number of years. We first calculate the time it took for each teacher to be promoted from one rank to another. Then we calculate the number of years for which, for each rank and education level in a county, half of the teachers were promoted within that time. Then for each rank and county we calculate the proportion of teachers promoted within that 'reasonable' time. We only include promotions before 2003. The ability index was constructed using the teachers' scores on an education test, a language test and a psychology test that were taken before the teacher began teaching. Scores are first standardized by the year in which they were taken, and then for each year the following measure is calculated: $[S - ave(S)]/ave(S)$ where S is the standardized score on the tests for a teacher, and $ave(S)$ is the average of the scores of the teachers in the same rank in the same school, to capture relative ability. The index was constructed using Principal Components Analysis (PCA). Dummies for the top and bottom 10% of this index are included. Evaluation score is measured on a scale of 1-4 (fail, pass, good, excellent).

C Appendix: Full Derivation of Equation (5)

To win a promotion, a teacher must place k^{th} or higher, or, must beat at least $n-k$ workers in pair-wise comparisons of performance. A teacher treats the effort of a competitor as a function of the random variable that is the competitor's skill, and treats all competitors as ex-ante identical in deriving the MPE. Given CDF $R(q_j)$ and PDF $r(q_j)$ over a competitor's performance, the probability that teacher i beats teacher g is $pr(q_i > q_g) = pr(e_i + s_i + \epsilon_i + \mu > e_g + s_g + \epsilon_g + \mu) = pr(e_g + s_g + \epsilon_g < e_i + s_i + \epsilon_i) = R(e_i + s_i + \epsilon_i)$. The probability that teacher i beats any opponent is $R(e_i + s_i + \epsilon_i)$ and the probability that teacher i loses to any opponent is $1 - R(e_i + s_i + \epsilon_i)$.

To finish exactly j^{th} from the top, a teacher must beat $n-j$ opponents and lose to $j-1$ opponents. The probability of doing so for any given partition of competitors, conditional on $i's$ luck is:

$$R(e_i + s_i + \epsilon_i)^{n-j}(1 - R(e_i + s_i + \epsilon_i))^{j-1} \tag{12}$$

The number of ways to chose $n-j$ elements from a collection of $n-1$ elements is $\binom{n-1}{n-j} = \frac{(n-1)!}{(n-j)!(j-1)!}$ so the probability of placing exactly j^{th} from the top conditional on ϵ_i is:

$$pr(place\ j^{th}\ out\ of\ n|\epsilon_i) = \binom{n-1}{n-j} R(e_i + s_i + \epsilon_i)^{n-j}(1 - R(e_i + s_i + \epsilon_i))^{j-1} \tag{13}$$

The probability of promotion conditional on ϵ_i is the sum of the conditional probabilities of placing first through k^{th}. Integrating out ϵ_i gives the unconditional probability of promotion:

$$p(s_i, e_i, \underline{e}) = \sum_{j=1}^{k} \binom{n-1}{n-j} \int R(e_i + s_i + \epsilon_i)^{n-j}(1 - R(e_i + s_i + \epsilon_i))^{j-1} f(\epsilon_i) d\epsilon_i \tag{14}$$

The order statistic part of this equation depends on the CDF $R(.)$ and PDF $r(.)$ of the random output of teacher $i's$ competitors. However, the only stochastic part of his output to integrate over (from his point of view) is ϵ, so $f(\epsilon)$ appears instead of an additional $r(.)$. Since e and s are constants, $d(e + s + \epsilon) = d\epsilon$.

To solve for dp/de, we must differentiate:

$$\frac{dp}{de} = \sum_{j=1}^{k} \binom{n-1}{n-j} \int (n-1)R(.)^{n-j-1}(1-R(.))^{j-1}r(.) - (j-1)R(.)^{n-j}(1-R(.))^{j-2}r(.)f(\epsilon)^2 d\epsilon \tag{15}$$

where $R(.) = R(e_i + s_i + \epsilon_i)$ and $r(.) = r(e_i + s_i + \epsilon_i)$. Since $(m-1)\binom{n-1}{n-m} = (n-(m-1))\binom{n-1}{n-(m-1)}$ and since the second expression is zero if $j = 1$, this becomes:

$$\begin{aligned} \frac{dp}{de} &= \sum_{j=1}^{k}(n-j)\binom{n-1}{n-j} \int R(.)^{n-j-1}(1-R(.))^{j-1}r(.)f(\epsilon)^2 d\epsilon \\ &- \sum_{m=2}^{k}(n-(m-1))\binom{n-1}{n-(m-1)} \int R(.)^{n-m}(1-R(.))^{m-2}r(.)f(\epsilon)^2 d\epsilon \end{aligned} \tag{16}$$

Letting $j = m - 1$, all the terms in the second sum cancel terms in the first sum, leaving only the term $j = k$:

$$\frac{dp}{de} = (n-k)\binom{n-1}{n-k} \int R(.)^{n-k-1}(1-R(.))^{k-1}r(.)f(\epsilon)^2 d\epsilon \tag{17}$$

We can simplify (17) by making use of change of variables $\Psi \equiv R(e + s + \epsilon)$:

$$\frac{dp}{de} = \int (n-k)\binom{n-1}{n-k} \Psi^{n-k-1}(1-\Psi)^{k-1} f(R^{-1}(\Psi) - s - e) d\Psi \tag{18}$$

D Appendix: Proof of Prediction 6

In years in which there are more than five years before a teacher is eligible for promotion, effort incentives are zero. That is, effort incentives are zero whenever $t \leq X - 5$. Choice of effort in year $t = X - 5$ is such that equation (6) is maximised, where $j = X - 5$:

$$EV_{X-5} = -c(e_{X-5}) + Ep_{X-4} \sum_{t=X-4}^{T} \beta^{t-(X-5)} U_h + (1 - Ep_{X-4})[Ep_{X-3} \sum_{t=X-3}^{T} \beta^{t-(X-5)} U_h$$
$$+ (1 - Ep_{X-3})\{Ep_{X-2} \sum_{t=X-2}^{T} \beta^{t-(X-5)} U_h + (1 - Ep_{X-2})(Ep_{X-1} \sum_{t=X-1}^{T} \beta^{t-(X-5)} U_h$$
$$+ (1 - Ep_{X-1}) Ep_X \sum_{t=X}^{T} \beta^{t-(X-5)} U_h)\}] \quad (19)$$

However, $Ep_{X-4} = Ep_{X-3} = Ep_{X-2} = Ep_{X-1} = Ep_X = 0$ since the teacher cannot be promoted until $t = X + 1$. As a result, all the terms in (19) drop out except for $-c(e_{X-5})$. Thus, effort is zero.

In the five years leading up to the year in which a teacher becomes eligible for promotion, effort incentives are increasing. That is, effort incentives are increasing from $t = X-4$ to $t = X$. We can write out expressions for choice of effort in years $j = X - 4$, $j = X - 3$, $j = X - 2$, $j = X - 1$, and $j = X$ using equation (6).

In $j = X - 4$:

$$EV_{X-4} = -c(e_{X-4}) + Ep_{X-3} \sum_{t=X-3}^{T} \beta^{t-(X-4)} U_h + (1 - Ep_{X-3})[Ep_{X-2} \sum_{t=X-2}^{T} \beta^{t-(X-4)} U_h$$
$$(1 - Ep_{X-2})\{Ep_{X-1} \sum_{t=X-1}^{T} \beta^{t-(X-4)} U_h + (1 - Ep_{X-1})(Ep_X \sum_{t=X}^{T} \beta^{t-(X-4)} U_h$$
$$+ (1 - Ep_X) Ep_{X+1} \sum_{t=X+1}^{T} \beta^{t-(X-4)} U_h)\}] \quad (20)$$

However, $Ep_{X-3} = Ep_{X-2} = Ep_{X-1} = Ep_X = 0$ since the teacher cannot be promoted until $t = X + 1$. As a result, (20) simplifies to:

$$EV_{X-4} = -c(e_{X-4}) + Ep_{X+1}\sum_{t=X+1}^{T}\beta^5 U_h \tag{21}$$

A teacher chooses effort in $j = X - 4$ in order to maximise:

$$\max_{e_{X-4}}\left\{Ep_{X+1}\sum_{t=X+1}^{T}\beta^5 U_h - c(e_{X-4})\right\} \tag{22}$$

The first order condition is:

$$c'(e_{X-4}) = \frac{dEp_{X+1}}{de_{X-4}}\sum_{t=X+1}^{T}\beta^5 U_h \tag{23}$$

This procedure can be repeated for $j = X - 3$, $j = X - 2$, $j = X - 1$, and $j = X$.

In $j = X-3$, after substiting $j = X-3$ and $Ep_{X-2} = Ep_{X-1} = Ep_X = 0$, the corresponding equation for expected lifetime utility is:

$$EV_{X-3} = -c(e_{X-3}) + Ep_{X+1}\sum_{t=X+1}^{T}\beta^4 U_h + (1 - Ep_{X+1})Ep_{X+2}\sum_{t=X+2}^{T}\beta^5 U_h \tag{24}$$

The first order condition is now:

$$\begin{aligned} c'(e_{X-3}) = {} & \frac{dEp_{X+1}}{de_{X-3}}\sum_{t=X+1}^{T}\beta^4 U_h + \frac{dEp_{X+2}}{de_{X-3}}\sum_{t=X+2}^{T}\beta^5 U_h \\ & - Ep_{X+2}\frac{dEp_{X+1}}{de_{X-3}}\sum_{t=X+2}^{T}\beta^5 U_h - Ep_{X+1}\frac{dEp_{X+2}}{de_{X-3}}\sum_{t=X+2}^{T}\beta^5 U_h \end{aligned} \tag{25}$$

Because of the second simplifying assumption made in the model, that performance of effort in any of the past five years before promotion has an equal impact on the probability of promotion ($dp_t/de_{t-1} = dp_t/de_{t-2} = ... = dp_t/de_{t-5}$), $\frac{dEp_{X+1}}{de_{X-3}} = \frac{dEp_{X+2}}{de_{X-3}}$ and the first order condition simplifies to:

$$c'(e_{X-3}) = \frac{dEp_{X+1}}{de_{X-3}}\left(\sum_{t=X+1}^{T}\beta^4 U_h + \sum_{t=X+2}^{T}\beta^5 U_h(1 - Ep_{X+1} - Ep_{X+2})\right) \tag{26}$$

Now we can compare (26) and (23). Again, because of simplifying assumption (2), we can substitute $\frac{dEp_{X+1}}{de_{X-3}} = \frac{dEp_{X+1}}{de_{X-4}}$ in (26) and compare:

$$\frac{dEp_{X+1}}{de_{X-4}} \sum_{t=X+1}^{T} \beta^5 U_h < \frac{dEp_{X+1}}{de_{X-4}} \left(\sum_{t=X+1}^{T} \beta^4 U_h + \sum_{t=X+2}^{T} \beta^5 U_h (1 - Ep_{X+1} - Ep_{X+2}) \right) \quad (27)$$

which holds as long as $Ep_{X+1} + Ep_{X+2} < 1$ since $\beta^5 < \beta^4$ and $\sum_{t=X+1}^{T} \beta^5 U_h < \sum_{t=X+2}^{T} \beta^5 U_h$. The sum of all expected probabilities is indeed less than one, since a teacher cannot be absolutely certain of promotion in the future.

In $t = X - 2$ the FOC simplifies to:

$$\begin{aligned} c'(e_{X-2}) = \frac{dEp_{X+1}}{de_{X-2}} [\sum_{t=X+1}^{T} \beta^3 U_h + \sum_{t=X+2}^{T} \beta^4 U_h (1 - Ep_{X+1} - Ep_{X+2}) + \\ (\sum_{t=X+3}^{T} \beta^5 U_h (1 - Ep_{X+1}(1 - Ep_{X+2} - Ep_{X+3}) - Ep_{X+2}(1 - Ep_{X+3}) - 2Ep_{X+3})] \end{aligned} \quad (28)$$

which again is larger than the FOC for $j = X - 3$. The same procdure can be followed for $j = X - 1$ and $j = X$.

In the years after a teacher becomes eligible for promotion ($j > X$), effort incentives are no longer increasing. Assume that $j > X$. We can write an expression for $j + 1$:

$$\begin{aligned} EV_{j+1} = -c(e_{j+1}) + Ep_{j+2} \sum_{t=j+2}^{T} \beta U_h + (1 - Ep_{j+2})[Ep_{j+3} \sum_{t=j+3}^{T} \beta^2 U_h + (1 - Ep_{j+3})\{Ep_{j+4} \\ \sum_{t=j+4}^{T} \beta^3 U_h + (1 - Ep_{j+4})(Ep_{j+5} \sum_{t=j+5}^{T} \beta^4 U_h + (1 - Ep_{j+5}) Ep_{j+6} \sum_{t=j+6}^{T} \beta^5 U_h)\}] \end{aligned} \quad (29)$$

and compare it to the expression for j:

$$EV_j = -c(e_j) + Ep_{j+1} \sum_{t=j+1}^{T} \beta U_h + (1 - Ep_{j+1})[Ep_{j+2} \sum_{t=j+2}^{T} \beta^2 U_h + (1 - Ep_{j+2})\{Ep_{j+3} \sum_{t=j+3}^{T} \beta^3 U_h + (1 - Ep_{j+3})(Ep_{j+4} \sum_{t=j+4}^{T} \beta^4 U_h + (1 - Ep_{j+4}) Ep_{j+5} \sum_{t=j+5}^{T} \beta^5 U_h)\}] \quad (30)$$

The discounting terms are all equivalent, the expected probability of promotion is non-zero in all years in both cases (since $j > X$), and there are an equivalent number of years in which current effort will affect a teacher's expected probability of promotion (five). As a result, effort incentives are not increasing.