

I am a development economist and my research examines labor and education market imperfections, especially around issues of asymmetric information and moral hazard. My research primarily relies on the use of large scale field experiments in partnership with firms, schools and governments. This allows me to causally identify the effect of various aspects of the principal-agent relationship, including altering contracts, incentives, monitoring, and information. In addition, I focus on measuring richer aspects of behavior, beliefs, preferences and information sets than would traditionally be available in administrative data using surveys, classroom observations and revealed preference measures with agents.

The research questions I study fall into two broad categories within a principal-agent framework i). understanding multi-tasking problems in teaching and ii). understanding the extent of private information of agents relative to the principal in labor, education and health markets.

Moral Hazard, Multi-tasking and the Development of Student Human Capital

A fundamental challenge in any principal-agent relationship is how best to incentivize agents where there are multiple outcomes of interest with varying levels of contractability due to differences in outcome observability, measurement or ability to incentivize. For example, in teaching, we want agents (teachers) to develop students' long run human capital, but can often only measure or incentivize short run outcomes like test scores. Depending on the underlying teaching production function, teacher preferences, incentives and cost of effort, incentivizing only short-run aspects of student human capital may actually be detrimental in the long run. To understand how best to incentivize, monitor and train teachers, we need a better understanding of how teacher effort impacts broader student skills, such as executive-functioning and socio-emotional skills, and the extent of crowd out when we focus on only an easily measurable subset of student outcomes.

My research address the following questions discussed in turn below:

- Is sustained attention a general skill that is transferable across domain and can be developed using high quality teaching pedagogy?
- Do test-score based incentives distort teacher effort toward certain actions which raise student test performance at the cost of non-test score skill development?
- What is the underlying teacher production function of teacher effort and pedagogy on student test-score and non-test score skill development?

Teachers may build human capital not only by teaching academic skills but by expanding the capacity for cognition. in **"Cognitive Endurance as Human Capital"** (QJE, 2025) with Supreet Kaur, Geeta Kingdon and Heather Schofield, we focus specifically on cognitive endurance: the ability to sustain effortful mental activity over a continuous stretch of time. As motivation, we document that globally and in the United States, the poor exhibit cognitive fatigue more quickly than the rich do across field settings; they also attend schools that offer fewer opportunities to practice thinking for continuous

stretches. Using a field experiment with 1,600 Indian primary school students, we randomly increase the amount of time students spend in sustained cognitive activity during the school day—using either math problems (mimicking good schooling) or nonacademic games (providing a pure test of our mechanism). Each approach markedly improves cognitive endurance: students show 22% less decline in performance over time when engaged in intellectual activities—listening comprehension, academic problems, or IQ tests. They also exhibit increased attentiveness in the classroom and score higher on psychological measures of sustained attention. Moreover, each treatment improves students' school performance by 0.09 standard deviations. This indicates that the experience of effortful thinking itself—even when devoid of any subject content—improves general cognitive capacity. Finally, we complement these results with quasi-experimental variation indicating that an additional year of schooling improves cognitive endurance, but only in higher-quality schools. Our findings suggest that teaching quality disparities may further disadvantage poor children by hampering the development of a core mental capacity.

A central challenge facing firms is how to incentivize employees. While objective incentives may be theoretically ideal, in practice they may lead employees to reduce effort on non-incentivized outcomes and may fail in settings where effort is weakly tied to outcomes. Subjective (manager-discretionary) incentives are hence the norm for the vast majority of employees. However, whether manager discretion is able to overcome this noise and distortion problem is unclear and may even perform worse than objective incentives if managers are inaccurate in their evaluations. In **“Subjective versus Objective Incentives and Employee Productivity”** (with Tahir Andrabi), we study the effect of subjective incentives (manager-discretionary performance evaluation) and objective incentives (test score-based) relative to no incentives for teachers using an RCT in over 200 Pakistani schools. First, we show that subjective and objective incentives both increase test scores and have similar magnitude effects. However, objective incentives decrease non-test score student outcomes relative to subjective incentives. Second, we show that teachers' effort responses are very different under each scheme, with attendance increasing under subjective and teaching quality decreasing under objective.

We then rationalize these effects through the lens of a moral hazard model with multi-tasking. We show teachers believe objective incentives are a noisier function of their effort than subjective incentives, and objective incentives distort effort toward narrow aspects of student learning. Then we use *within-treatment* variation to isolate the causal effect of contract noise and distortion. We show that noisier incentive schemes lead to a smaller effort response and distorted incentives shift teacher effort to certain actions. Combined, differences in the perceived noise and distortion of subjective versus objective incentives can explain a majority of the difference in reduced form effects between the two contract types.

Finally, in ongoing work, **“Estimating the Production Function of Teaching”** with Sabrin Beg, James Berry, Anne Fitzpatrick, and Adrienne Lucas, we use the staggered roll-out of an intervention focused on giving teachers feedback on a one of 12 different dimensions of pedagogy (e.g. time on task, classroom climate, quality of feedback, etc) in 20,000 primary schools across Andhra Pradesh, India. We use exogenous variation

in which dimension of teaching a given school cluster was told to focus on. We then measured teacher behavior from classroom observations and thrice-yearly student standardized student assessments to estimate the returns to teacher behaviors on student learning. We can also see the extent which certain teacher behaviors improve content knowledge at the cost of more well-rounded student skills.

Information and Contracting Frictions

The second closely related line of work focuses on understanding the information principals versus agents have about the underlying production function or the agent's type. "Type" here includes the traditional notion of ability type, but also other aspects such as the extent of present bias or knowledge a teacher has about their student's baseline knowledge. My work explores how we can measure this private information "gap" (how much information principals versus agents have) and how to design contracts and incentives given the extent of this gap. Within this area, I explore a number of research questions in this areas including:

- Does offering performance pay allow schools to attract and retain better teachers as well as those more responsive to incentive pay? How much private information to teachers have about their own quality beyond what employers know?
- How do financial stakes and better information about worker effort impact the extent of gender discrimination in performance evaluations?
- How much of the social network hiring premium is due to information frictions versus production complementarities?
- When is it better to provide teachers discretion in what topics to teach? How does this vary based on the extent of teacher knowledge, preferences and cost of effort?

Attracting and retaining high-quality teachers has large social benefit, but schools often struggle to identify good teachers ex-ante. In **"Inducing Positive Selection through Performance Pay: Experimental Evidence from Pakistani Schools"** with Tahir Andrabi, we use teachers' contract choices and a randomized controlled trial of performance pay with 7,000 teachers in Pakistan to study whether performance pay affects the composition of teachers. Consistent with adverse selection models, we find that performance pay induces positive sorting: both among teachers with higher latent ability and among those with a more elastic effort response to incentives. Using two additional treatments, we show effects are more pronounced among teachers with more private information about their own quality and teachers with lower job switching costs. Accounting for these sorting effects, the total effect of performance pay on test scores is twice as large as the behavioral effort response from the existing stock of teachers, suggesting that analyses that ignore sorting effects may substantially understate the effect of performance pay.

Pakistan ranks in the lowest decile in female labor force participation, and even in sectors where women are more prevalent, such as teaching, they earn 70 cents for each dollar

men earn. While we have extensive evidence on the prevalence of gender bias in hiring, promotions and wages, we know less about the mechanisms underlying this bias and the extent to which certain personnel policies may mitigate or exacerbate these biases. In **“Understanding Gender Discrimination by Managers”**, I conduct a large scale field experiment with 3,600 employees in 250 schools and randomly vary i). how often managers observe a given employee and ii). whether manager evaluations affect employee’s pay or are just used for feedback. First, I find when there are no financial stakes associated with performance evaluations, there is minimal difference in scores between men and women. This holds even when controlling for a rich set of controls of teacher productivity, such as value-added, clock in and out time, time use, and pedagogy measured via classroom observations.

In contrast, when principals’ evaluations determine teachers’ end of year raise, we see that female teachers receive 20% lower raises, controlling for productivity. However, when principals are randomly assigned to conduct more frequent classroom observations of the teacher, this increases the evaluation of female teachers and closes two-thirds of the gender gap under financial stakes. To understand mechanisms, I conduct a follow up vignette survey to test whether our results are due to differential manager expectations about employee reactions to low raises (e.g. higher turnover by men) or differential perceived deservedness of scarce financial resources. The results favor this second explanation as managers favoring single-earner (lower income) households over dual-earner (higher income) households, which is highly correlated with employee gender. Combined this suggests that improving the accuracy of manager information could close the gender gap in performance evaluations, even in high stakes settings.

Network-based hiring dominates labor markets in developing countries, yet whether this reflects efficient screening or market failures remains unclear. In **“Social Networks and Hiring: Experimental Evidence from Pakistan”** with Maryiam Haroon, we investigate two mechanisms: information asymmetries favor workers with known productivity, or connections generate match-specific returns through productivity complementarities, enhanced trust, or reduced monitoring costs. Through a randomized controlled trial with 87 construction managers and 1,500 workers in Pakistan, we provide objective productivity information, introduce performance-based pay, and randomly remove managers from worksites. We find networks operate through search friction reduction rather than productivity selection—connected workers exhibit no productivity advantage yet secure employment 30% faster. Providing productivity information generates substantial wage gains (\$13.95 monthly income, \$0.40 daily wages), but exclusively for connected workers who simultaneously reduce hours worked. Unconnected workers experience null effects despite identical information, revealing that productivity signals complement rather than substitute for social capital. Performance incentives fail to shift hiring patterns or productivity. Exploiting variation in manager presence, we find connected workers face hiring advantages only when their manager is present, yet productivity remains unchanged—indicating networks provide non-transferable private information about reliability rather than productivity complementarities.

In ongoing work, **“Structure vs. Discretion: Experimental Evidence on the Effects of Teacher Autonomy”** with Tahir Andrabi, Asim Khwaja and Isabel MacDonald, we seek to understand when is better to give teachers more or less discretion to address diverse learning levels within the classroom. Education systems around the world face a “learning crisis;; of low average learning and high learning variation with each single classroom. The COVID-19 pandemic has exacerbated these problems, and threatens lasting consequences for a generation of students. Given low overall teaching quality and the challenge teachers face of trying to adjust lessons to a variety of different learning levels, numerous studies have suggested that giving teachers detailed lesson plans can help improve instructional quality. However, teachers may also have substantial private information about what lessons are likely to be most effective with certain students.

To test for whom (if anyone) discretion works best, we partner with a public school system with 500 schools and 60,000 students to randomly assign teachers to (i) a control group (business as usual), (ii) a high-discretion treatment, where they received information on student learning levels and materials to tailor instruction, or (iii) a low-discretion treatment, which added structured daily guidance on student-level use of the materials. Preliminary results show that the intervention led to gains of 0.14–0.17 standard deviations in student learning outcomes. High-discretion only outperformed low-discretion for teachers with extremely high locus of control and low cost of effort.

We follow up this study by tracking teachers and students two years later to measure learning and teacher behavior in **“Persistence and Behavior Change in Public Schools”** with Tahir Andrabi and Asim Khwaja. The project evaluates the persistence and scalability of the program by (i) examining whether initial gains in student learning sustain after direct support ends, (ii) testing whether teachers continue to adopt improved instructional practices with new student cohorts, and (iii) assessing whether repeated use of the intervention improves, maintains, or diminishes its effectiveness. Findings from this study will offer evidence on the durability of low-cost, information-driven interventions for improving foundational learning and highlight strategies for institutionalizing such practices in public education systems.

In **“Effects of Classroom Observation and Feedback on Teacher Effort and Student Learning”** with Sabrin Beg, James Berry, Anne Fitzpatrick, and Adrienne Lucas we seek to understand whether a light touch intervention where teachers are provided feedback about their teaching quality can impact teacher behavior and student learning. In particular, we seek to understand under what conditions do teachers believe the feedback provided and act on this information. We vary the agent providing the information and extent to which there are longer term reputational consequences. We randomize treatment across 600 schools and measure teacher beliefs, behavior and student learning.

Finally, in **“Optimal Incentives for Vaccination under Present Bias and Uncertainty”** with Rachel Glennerster and Maryiam Haroon, we leverage an at-scale incentive for immunization programs in Pakistan to understand whether it is possible to use prior behavior or self-selection to identify who will benefit from commitment devices. House-

holds received baseline payments and were randomly assigned to additional incentives tied to strict, moderate, or no deadlines. The design tests whether tailoring incentive structures could potentially have large benefits for timely vaccination and if agents have private information about their extent of present bias compared to what can be predicted from past behavior.