ABSTRACT

Title of dissertation:	MANAGERIAL INCENTIVES IN PUBLIC SERVICE DELIVERY: EVIDENCE FROM SCHOOL-BASED NUTRITION PROGRAMS IN RURAL CHINA
	Sean Sylvia, PhD, 2014
Dissertation directed by:	Professor Vivian Hoffmann Department of Agricultural and Resource Economics

Growing evidence indicates that weak or misaligned incentives facing providers pose a significant barrier to service delivery in many developing countries. To address weak supply-side incentives, performance pay and related approaches explicitly linking provider pay and performance have become increasingly common in public service delivery. Despite the growing prominence of these approaches, however, many conceptual issues surrounding the use of performance pay in this context remain unexplored. A fundamental one is the role of performance pay for managers in the organizations commonly tasked with delivering public services. Although a large literature examines performance pay for managers in private firms, much less is known about the use of performance pay for their counterparts in public service organizations. Improving public service delivery may nonetheless depend heavily on aligning the incentives of managers with social objectives.

Drawing on a large-scale field experiment involving 300 primary schools in rural China, this dissertation explores how performance incentives for school administrators affect their implementation of new, school-based nutrition programs targeting anemia. School-based nutrition programs are an important function of schools, particularly in settings with less developed public health infrastructures. Weak incentives for schools to effectively implement these programs are compounded

as these programs compete with more traditional functions for finite school resources. I report the results of this field experiment which was designed to test three main issues concerning the use of performance incentives for school administrators in this context.

First, I study the effect of offering administrators performance pay contracts tied to reductions in school-level anemia prevalence. As part of the experiment, a subset of schools were randomly allocated to receive one of two levels of performance incentives for reductions in student anemia or to a no-incentive comparison group. I find that large incentives led to meaningful reductions while smaller incentives (10% of the size) were ineffective in reducing anemia. Further, I find that an important channel through which large incentives impacted student nutrition was by motivating administrators to engage households and influence feeding at home. I discuss the implications of this finding for the design of performance incentives tied to jointly produced outcomes.

Second, I study the impact of providing administrators with more resources to implement a nutrition program and how this interacts with performance incentives. To test this, schools were orthogonally assigned to two levels of block grants within each level of performance incentives. I find that, absent explicit anemia-based incentives, increasing the size of block grants under the control of administrators led to sizable reductions in anemia prevalence but were nearly twice as costly as performance incentives. This impact was not purely the result of additional inputs; larger block grants also caused a more efficient use of inputs and an increase in effort devoted to reducing anemia. I also find that additional resources and incentives are substitutes in this context. I provide evidence that this substitution is due, at least in part, to incentives re-framing the task of implementing the nutrition programs from one that was part of the professional role of administrators to one that was not.

Finally, I approach the health promotion and education roles of schools as a multi-tasking problem and use remaining experimental groups to examine how performance incentives for school administrators to reduce anemia and improve test scores each affect anemia prevalence and academic performance. Although the theory of multitasking is well-developed, there are few empirical studies testing this theory directly. I emphasize three main findings. First, incentives in the two dimensions (given in the context of an anemia reduction program) both led to significant reductions in anemia prevalence. Second, anemia-based and test-based incentives serve as substitutes in the direction of anemia reduction: providing administrators with both types of incentives did not lead to significantly larger reductions in anemia. Third, I find that anemia incentives caused an allocation of resources away from education 'inputs' but this did not lead to significantly lower student performance on standardized exams after one year. These results reflect that test-based incentives are well-aligned with improving nutrition, but anemia-based incentives are not well aligned with effort to improve academic performance. Strengthening incentives to improve academic performance while also emphasizing the relationship between good nutrition and academic performance may therefore be sufficient to motivate administrators to effectively implement school-based anemia reduction programs while causing less reallocation of resources away from education.

MANAGERIAL INCENTIVES IN PUBLIC SERVICE DELIVERY: EVIDENCE FROM SCHOOL-BASED NUTRITION PROGRAMS IN RURAL CHINA

by

Sean Yuji Sylvia

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Vivian Hoffmann, Chair
Professor Scott Rozelle
Professor Grant Miller
Professor Kenneth Leonard
Professor Anna Alberini
Professor Peter Murrell

Dedication

To Nana, Zoe and Owen.

Acknowledgments

This research was funded by grants from the National Institutes of Health (Grant Number: R01 HL106023) and the Walter H. Shorenstein Asia-Pacific Research Center at Stanford University. I also acknowledge generous support while working on this dissertation from the Stanford Center for International Development (SCID).

There are many individuals to whom I owe a great dept of gratitude. First and foremost, I thank my family. I thank my wife, Na, for her support, understanding, encouragement and love. Without her my life would be much less wonderful. My parents encouraged me to explore and provided a safety net that allowed me to do so. My grandparents, Jiichan and Baachan, taught me to question everything and made the world a small place.

I am extremely fortunate to have many individuals that I call a mentor. It is difficult to express my gratitude to Shi Yaojiang and Scott Rozelle, two individuals who have shaped me academically and personally more than they can imagine. I am indebted to both for their guidance and friendship. Shi laoshi showed me what a fascinating and glorious place China is. Scott took me in as his own, taught me how to be a student of China, and how to be a passionate academic. 一日为师，终身为父。 I also thank my "real" advisor, Vivian Hoffmann, for her support and guidance navigating the maze of graduate student-dom; and Grant Miller for his guidance on my dissertation and life after graduate school.

The fieldwork for this dissertation would not have been possible without the dedication of the REAP team in China: the "heroes" of CCAP and the NSDRC. I owe a special thanks to Luo Renfu for his hard work and comradeship during my fieldwork; also to Zhang Linxiu for giving me an academic home in China and being an example of commanding elegance.

I am also grateful to Grant Miller, Marcos Vera-Hernández, Scott Rozelle and Luo Renfu for allowing me to include our joint work in my dissertation.

Finally, I want to acknowledge the support of the great friends and colleagues: Anand Murugesan, Dana Anderson, Ariel Ortiz Bobea and my other AREC classmates; Alexis Medina, Prashant Loyalka, Matt Boswell, James Chu, Ma Xiaochen, Mo Di and the rest of the REAP team at Stanford.

Contents

List of Tables

List of Figures

Chapter 1: Introduction

Much research in development economics has concerned the question of how best to deliver public services to the poor. Focusing on the provision of education and healthcare, this research has considered how public services should be provided so to most cost-effectively improve population outcomes. A large portion of the research on this issue to date has focused on addressing the many demand-side barriers that prevent individuals from investing in education or seeking healthcare optimally. Equally daunting barriers, however, exist on the supply-side with providers. A growing literature documents the poor provision of services in many developing countries (World Bank 2004; Banerjee et al. 2004; Das et al. 2008; Berendes et al. 2011; Sylvia et al 2013). Although the underlying reasons are complex and incompletely understood, the culprit is not simply lack of resources, inadequate training or deficiencies in provider knowledge. Supply-side incentives are also often poorly aligned with social objectives.

Perhaps the most striking example is pervasive absenteeism among education and health workers in many parts of the world (Chaudhury & Hammer 2003; Banerjee, Deaton and Duflo (2004); Kremer et al. 2005; Chaudhury et al. 2006; Lewis 2006; Banerjee & Duflo 2006; Das, Hammer and Leonard 2008; Berendes et al. 2011). When present, providers also often fail to do what is within their knowledge and means. Several studies, for instance, have documented a substantial "know-do gap" (or failure to do in practice what they know to do in principle) among clinicians (Das and Hammer 2004; Alcazar et al. 2006; Chaudhury et al. 2006; Das & Hammer 2007; Leonard & Masatu 2010; Das et al. 2012). Educator and healthcare provider effort may also not be well-aligned with improving education and health outcomes. For example, inappropriate incentive structures can lead healthcare providers to provide treatment that is unnecessary and potentially harmful (the over-prescription of antibiotics in China in response to distroted drug price schedules is a clear example – Currie et al. 2011, Yip et al. 2014).

To better align provider incentives with social objectives, "results-based financing" approaches have become increasingly common in public sector service delivery (Oxman & Fretheim 2008; Eichler & Levine 2009). Drawing on the logic of performance pay in human resource management (Lazear 1995; Hall & Liebman 1998; Lazear 200), these approaches provide direct financial rewards for achieving pre-specified performance targets.

While current evidence suggests that performance pay and related approaches hold promise as a way to improve education and health outcomes, many conceptual issues about the appropriate use of performance pay in public service delivery remain unexplored. A fundamental one is the role of performance pay for managers in the organizations that most commonly deliver these services. Although executive and managerial performance pay in the private sector is the subject of a large literature, much less is known about performance pay for public-sector managers.[1] Further, existing studies on performance pay in the public sector focus largely on incentives for front-line service providers, including teachers and health workers (Rasul & Rogger 2013).[2] Incentives for managers have broader aims, however – including encouraging the desired allocation of resources under managerial control (Holmstrom & Costa 1986). Although evidence on the incentives of public sector managers is limited, private sector managers exert substantial influence on overall organizational performance (Bloom et al. 2011; Bandiera et al. 2007). In public sector organizations, managerial incentives may play a role in improving allocative efficiency within the organization and encouraging innovation in service delivery.[3] On the other hand, high-powered incentives may be ineffective or lead to undesired consequences due to distinct features of public service (Dixit 2002).

[1] The literature on managerial incentives in the private sector includes, for example, Jensen and Murphy (1990); Hall and Liebman (1998); Hall and Murphy (2003); and Oyer and Shaefer (2005). See Murphy (1999) for a review of CEO pay.

[2] Behrman et al. (2013) included rewards for school administrators. These rewards, however, were bundled with a larger intervention that rewarded students and teachers as well and compared to student-only and teacher-only interventions. This design precluded the isolation of the effect of school administrator awards alone.

[3] Although I focus on delivery by the public sector, the complications of providing public services may have less to do with the ownership of organizations that provide them and more to do with the nature of the services they provide (Besley & Ghatak 2007; Dixit 2012). Services provided by public agencies are often complex and the goals of organizations multi-faceted as a result (Dixit 2002). In addition, they are subject to well-known market failures that lead to under-provision (in the health sector, for example, patients are unable to judge the quality of care they recieve – Arrow 1963). Nevertheless, "remediable" inefficiencies (in the words of Williamson, 1996) may be particularly pernicious in the public sector.

In this dissertation, I study the role of managerial incentives in service delivery by analyzing a large-scale policy experiment in rural China. My specific focus is school administrators during the introduction of a new school-based nutrition program targeting iron deficiency, the most common micronutrient deficiency affecting school-aged children in China (Press Office of the Ministry of Health 2012). Although school-based health and nutrition programs are an important function of schools (particularly in developing countries) (Bundy and Guyatt, 1996; Jukes, Drake & Bundy, 2008; Orazem, Glewwe & Patrinos, 2008), weak incentives for schools to focus on health improvement may limit the effectiveness of these programs. Incentives facing educators in developing countries are often weak in general (cf.World Bank 2004; Duflo & Hanna 2005; Chaudhury et al., 2006; Banerjee & Duflo 2006). Further, even motivated educators may focus on traditional responsibilities over health promotion. Although health promotion and education may be complementary functions, they compete for finite school resources. Poor educator incentives for improving health may therefore reduce the ability of school-based health and nutrition interventions to improve student health outcomes through reduced compliance or diversion of resources to more traditional functions.

This context provides an interesting case study of managerial incentives for several reasons. First, schools in rural China—and primary school administrators (principals), in particular—are legally responsible for the wellbeing of their students (Anonymous 1995). Second, school administrators are often given a good deal of de facto discretion in their management of the school, though this varies to some degree by locality. Third, school-based health programs are a prime example of the most common type of change among public agencies—the addition of new programs on top of existing responsibilities, scenarios in which managerial incentives are central to success (Wilson 1989). Wilson (1989) and others emphasize that public sector agencies tend to resist new tasks if they do not fit well into the organization's self-defined mission (Wilson 1989; Dixit 2002). In this case, school administrators may resist a new program to reduce anemia if they do not view it as complementary to their core responsibility of educating students.

Within this context, I study several features of performance incentives for school administrators. First, I study the impact of administrator performance incentives tied to student nutrition, rewarding administrators based on school level changes in the rate of anemic students over the school year. I evaluate impacts on how the school-based anemia reduction program was implemented and on anemia prevalence. I also study how the effectiveness of this performance incentive varies with the size (or strength of the incentive). Specifically, I compare small incentives with large incentives ten times the strength. Previous studies from a variety of contexts have shown small incentives or prices leading to large changes in behavior (Kremer & Miguel 2007; Thornton 2008; Banerjee et al. 2010; Cohen & Dupas 2010; Karlan et al. 2011; Duflo et al. 2011). The findings in these studies suggest that the existence of incentives or a price may be more important than the actual level of incentives. My aim is to test this directly in the context of performance incentives.

Second, because an important distinction between managers and front-line workers is the control they exert over use of an organization's resources (and it is common for performance pay programs to be introduced in conjunction with increases in resources or training - Oxman 2008), I study how the responses of managers to performance incentives vary with budget size.[4] Specifically, I test whether performance incentives and the amount of resources under control of administrators are complements or substitutes. Complementarity is possible for intuitive reasons. For example, managers may wish to respond to performance incentives, but resource (or decision-making) constraints hinder their ability to do so, limiting the circumstances under which performance pay will be effective. There are also theoretical rationales for substitution, however. One possibility, for instance, is that if the marginal return to managerial effort declines with resources allocated to a task (and effort is costly), managers with larger budgets may simply substitute organizational resources for their effort in the production of contracted outcomes. How incentives and budgetary resources interact may also be less clear if a larger budget (and more discretion over the use of funds) itself has a motivational effect on school administrators. The public administration literature emphasizes this possiblity if an increase in the amount of resources under the control of school

[4]This section and the first include material from joint work with Renfu Luo, Grant Miller, Scott Rozelle and Marcos Vera-Hernandez.

administrators is also seen as an increase in autonomy (Rose-Ackerman 1986; Perry & Wise 1990).

Third, I approach these two roles of schools as a multi-tasking problem and examine how provider (school administrator) incentives to reduce anemia and improve test scores each affect anemia prevalence and academic performance. Further, I test whether these two incentives are complements or substitutes. Although the theory of multitasking (following from the work of Holmstrom and Milgrom 1991) is well developed, there are few empirical studies testing this theory directly. Further, existing studies focus on multitasking in the context of rewarding an imperfect measure of an unobservable outcome that is of ultimate value to a principal (for example, rewarding quantity when quality is also valued - Hong et al. 2013). In this study, I abstract from the issue of observability and focus on the effects of incentives for outcomes (potentially) requiring different inputs. Specifically, I aim to test how strengthening incentives in one direction affects the incented outcome as well as a second outcome that is also valued. I also ask whether there is a benefit in terms of either outcome gained by attempting to deal with multitasking by strengthening incentives in both directions simultaneously. School-based nutrition programs provide an ideal setting for these tests as improving nutrition and academic outcomes are not completely orthogonal (as is generally the case for tasks assigned to a single organization or agency). Further, the relationship between the two roles is asymetric (nutrition is an input to academic performance but academic performance is not an input to nutrition) which allows me to examine how incentives affect more or less well-aligned outcomes. The issues addressed here have important implications for the deisgn of performance incentive schemes for health delivery, in particular, given that these often base rewards on multiple outcome measures produced by more or less complementary inputs (Sherry, Bauhoff and Mohanan 2013).

By focusing on school managers (administrators), this dissertation contributes to a small but growing literature on the use of performance pay in the delivery of public services in developing countries. Several recent studies have examined the effects of performance pay structured as personal income for front line providers of health and education services (Lavy 2002; Lavy 2009; Glewwe et al. 2010; Singh 2011; Muralidharan & Sundararaman 2011; Duflo et al. 2012; Ashraf, Bandiera and Jack 2013; Behrman et al. 2013). While the effort of front line workers may ultimately have a significant influence on overall organizational efficiency, they generally have limited control over organizational resources (including the effort of others in the organization). Their ability to fully respond to outcome based incentives may therefore also be limited since they are less able to draw on specific knowledge and innovate in service delivery. On the other hand, performance incentives at the institutional level (for NGOs or local governments, for example) are more likely to influence resource allocation and improve allocative efficiency. Such institutional level incentives have also been studied in a number of contexts (Basinga et al. 2011; Gertler and Vermeersch 2012; Olken, Onishi and Wang 2012; Yip et al. 2014). Compared to individuals, organizations may also be less subject to behavior pitfalls commonly referenced in the literature on performance incentives (namely that extrinsic rewards can crowd out intrinsic motivation – Frey and Jegen 2001, Gneezy et al. 2011, Kamenca 2012). Performance incentives targeting organizations, however, are commonly structured as potential increases in budget revenue and it is unclear to what degree this motivates individuals within organizations (and how). Incentives for managers (in the hierarchical organizations that most commonly provide public services in developing countries - WDR 2004) may be an important middle ground.

The rest of the dissertation is organized as follows. The next chapter reviews the recent literature on incentives in public service delivery. Chapter 3 provides more background information on iron deficiency and its prevalence in China. Chapter 4 discusses the conceptual framework and research questions in more detail. Chapter 5 presents the experimental design, data collection and empirical approach. Chapters 6 presents the experimental results. The final chapter concludes.

Chapter 2: Performance Pay and Managers in Public Service Delivery

Eichler (2006) defines pay-for-performance as "a transfer of money or material goods conditional on taking a measurable action or achieving a predetermined performance target." [1] The intuition underlying performance pay and related approaches is the hidden action principal-agent problem in which a principal wishes to better align the incentives facing an agent with her own utility function. Performance pay approaches attempt to accomplish this by providing rewards (contracts) based on some performance measure thought to reflect effort that is valuable to the principal. Such performance incentive contracts have, of course, long been used to solve hidden action principal-agent problems in a wide range of contexts. Examples in private firms include sales commissions or CEO pay tied to firm value (see Lazear 1995; Hall & Liebman 1998; Lazear 2000).

Amid evidence of weak and misaligned incentives among those providing public services, performance pay approaches have become an increasingly common feature of service provision in wealthy and developing countries alike.[2] In 2011, 28 of the 34 OECD countries had introduced some from of performance pay in the public sector (Hasnain, Manning and Pierskalla 2014). Prominent examples in wealthy countries include the British National Health Services quality and outcomes framework and the introduction of teacher performance pay in a number of states in the U.S.. Performance pay is also becoming more prominent in developing countries. A number of international aid organizations, for example, have introduced performance pay under the umbrella of "results-based financing." Governments are also increasingly experimenting with performance

[1]Several terms are used in the literature for performance incentive contracts including performance pay, pay-for-performance, results-based financing, performance-based funding and output-based aid.

[2]On one level, performance pay is no different than user fees. The distinction is that government funding of public services is traditionally through set operating budgets allocated by higher levels. These are commonly based on volume of services that are provided by an organization (i.e. capitation) and not related to performance measures.

pay approaches. Rwanda, beginning in 2005, began to implement a national P4P scheme that supplemented health centers' input based budgets with payments tied to 14 maternal and child healthcare output indicators (Basinga et al. 2011). In 2009, China introduced a new 'Teacher Performance Pay Policy', stipulating that a portion of teacher pay increases be tied to performance evaluations.

Despite increasing interest among policy makers in using performance pay as a way of improving public sector productivity, empirical evidence on the effectiveness of performance pay approaches in the delivery of public services remains limited. A recent Cochorane report on pay for performance in developing country health delivery, for instance, concluded that evidence was too week to draw conclusions (Witter et al. 2012). Although evidence is growing, additional evidence is needed as these approaches are increasingly adopted. Evidence from public service organizations is particularly important given often-voiced concerns about the use of high-powered incentives in the public sector (Muralidharan 2012). These concerns include, for example, the difficulty of appropriately designing rewards given the complex nature of services provided by public service organizations and lack of clearly defined goals (specifically in the presence of multitasking and multiple principals – Dixit 2002) and the potential for extrinsic incentives to crowd out intrinsic motivation or degrade professional norms (Frey and Jegen 2001; Leonard and Masatu 2010).

This chapter outlines some of the recent evidence on the use of performance pay in education and health sectors in developing countries. The design of performance pay schemes varies significantly in in a number of dimensions. In my discussion, I focus on the dimension of who is rewarded as this is most illustrative of the potential role of incentives for managers. In general, studies to date have focused on either incentives for front line workers or incentives for organizations. The final section discusses some relevant evidence on managers and their incentives.

2.1 Rewarding Individuals

Studies of incentives as personal income to provide public services in developing countries have focused on front line staff such as teachers and health workers. Teacher performance pay, for instance, has now been studied in a number of developing countries (Lavy 2002; Lavy 2009; Glewwe et al. 2010; Muralidharan & Sundararaman 2011; Duflo et al. 2012; Behrman et al. 2013). Some of the best recent evidence on teacher incentives in developing countries comes from a recent large-scale randomized trial among primary school teachers in the Indian state of Andhra Pradesh (Muralidharan & Sundararaman, 2011). After 2 years of the program, incentives led to achievement gains of 0.27 standard deviations in math and 0.17 standard deviations in language. After five years, students who attended all five years of primary school under the teacher incentive program scored 0.54 standard deviations higher in math and 0.35 standard deviation higher on language tests compared to students in control schools. Students also scored significantly higher on subjects that were not incentivized under the program. Further, the authors test extensively for perverse outcomes such as "teaching to the test", reduced attention to non-incentivize subjects and teachers attempting to game the contract and find no evidence that any of this occurred.

Other studies of teacher incentives, however, have shown more mixed results. Reporting the results of a randomized trial in Kenya that provided group incentives to teachers based on test scores, Glewwe et al. (2010) found that, although test scores increased in the short run in incentive schools by 0.14 standard deviations, these gains faded after the program ended. Further, they find that, while there was no effect on teacher attendance or homework assignment, there was an increase in test preparation sessions. They interpret these results as evidence that teachers increased effort on test preparation in incentivized subjects but not on effort that contributed to longer term learning. In a more recent experiment in Mexican high schools, Behrman et al. (2013) find no impact of incentives for mathematics teachers.

Singh (2011) and Ashraf, Bandiera and Jack (2013) provide sobering results for rewarding individuals to provide health services in developing countries. Singh (2011) studies rewards for Indian day care workers tied to reductions in malnutrition. Performance incentives for day care workers alone were ineffective, but combining incentives with nutrition information to mothers led to a decline in malnutrition of 4.2% after three months. Ashraf, Bandiera and Jack (2013) find that financial incentives for hairdressers and barbers in Zambia to sell condoms in their shops was ineffective. Non-financial incentives that provided social recognition, however, led to significant gains in condom sales.

2.2 Rewarding Organizations

While the effort of front line workers may ultimately have a significant influence on overall organizational efficiency, they generally have limited control over organizational resources (including the effort of others in the organization). Their ability to fully respond to outcome based incentives may therefore also be limited since they are less able to draw on specific knowledge and innovate in service delivery (one of the main arguments of performance pay proponents). On the other hand, performance incentives at the institutional level (for NGOs or local governments, for example) are more likely to influence resource allocation and improve allocative efficiency.

Such institutional level incentives have also now been studied in a number of contexts (Basinga et al. 2011; Gertler and Vermeersch 2012; Olken, Onishi and Wang 2012; Yip et al. 2014). For the most part, these studies have focused on organizations providing health services. [3] Basinga et al (2011) and Gertler and Vermeerch (2012) evaluate the National P4P scheme in Rwanda which supplemented health centers' input based budgets with payments tied to 14 maternal and child healthcare output indicators. Using a randomized roll-out among a group of 166 health care facilities, these studies find that the program significantly increased a number of incentivized measures including institutional deliveries and preventative health service use by children under 59 months. The program also had a large effect on the weight-for-age of children 0-11 months and height-for-age among children 24-49 months and significantly reduced the "know-do" gap (the difference between what providers know to do and what they do in practice). The researchers also find evidence of strong complementarity between performance incentives and provider skill.

Olken, Onishi and Wong (2012) study a program in Indonesia that included incentives for local governments to provide health and education services. Under the program, villages were given block grants to provide schooling and maternal and child health services. In a randomly

[3] For reviews of the use of performance pay in health, see Eldridge and Palmer 2009, Witter et al. 2012, and Miller & Babiarz 2013.

chosen subset of villages, subsequent year block grants were tied to performance on 12 maternal and child health and education indicators relative to other villages in their region. Incentives led to better performance on health indicators but had no impact on education. The authors estimate that 50-75% of the total impact of block grants on health indicators was attributable to incentives.

Yip et al. (2014) conducted a randomized evaluation of pay-for-performance with capitation grants (against the status quo of fee-for-service) for rural township health centers and their subordinate village clinics in Ningxia, a relatively poor province in northwest China. Under the intervention, funding for township health centers was switched from a fee-for service basis to a capitation grant. At the beginning of the year, treatment townships received 70% of this grant. The remaining budget was then dispersed based on performance assessments tied to antibiotic prescription rates and patient satisfaction (and were subject to a quantity threshold). The authors found that pay-for-performance with capitation led to a 15% drop in rates of antibiotic prescription but found no impact on other outcomes (patient satisfaction, volume of outpatient visits and health spending).

Compared to individuals, organizations may also be less subject to behavior pitfalls commonly referenced in the literature on performance incentives (namely that extrinsic rewards can crowd out intrinsic motivation – Frey and Jegen 2001, Gneezy et al. 2011, Kamenca 2012). Performance incentives targeting organizations, however, are commonly structured as potential increases in budget revenue and it is unclear to what degree this motivates individuals within organizations (and how). Organizations still need to solve internal principal agent problems with front line workers. For instance, the contracts for health centers in Rwanda studied in Basinga et al. (2011) and Gertler and Vermeersch (2012) paid facilities, not individuals, based on the provision of a basket of services, however facilities could use the p4p funds at their discretion and more than 77% was used for staff compensation. Of course, incentives could be structured as personal pay tied to aggregate performance. For lower-level workers, however, such a reward structure could create incentives for free riding (Holmstrom 1982). Consistent with this, Muralidharan & Sundararaman (2011) find that group incentives for teachers consistently under perform individual incentives.

2.3 Managers

Incentives for managers (in the hierarchical organizations that most commonly provide public services in developing countries - WDR 2004) may functionally be an important middle ground between incentives for front line workers and organizations as a whole. The role of performance incentives for managers in the public sector, in both wealthy and developing countries, has not been well studied, however. Although executive and managerial performance pay in the private sector is the subject of a large literature[4], much less is known about performance pay for public-sector managers.

There is some evidence on the correlation between incentives and performance for public sector managers from two recent observational studies. Binderkrantz and Christiansen (2012) study performance pay reform in Denmark that included financial rewards in executive contracts tied to achievement of performance targets. Using data on 124 agency heads, they find little correlation between bonus payments and performance indicators. Rasul & Rogger (2013) study incentives for bureaucrats in the Nigerian Civil Service using a unique dataset on 4700 projects across 63 civil service organizations responsible for these projects. They estimate that a one standard deviation increase in performance incentives is correlated with a 14% *drop* in project completion rates. Further, this negative correlation is strongest for more complex projects and projects that organizations implement less frequently. Importantly, Rasul & Roger (2013) also document that a one standard deviation increase in autonomy corresponds to an 18% increase in project completion among Nigerian Bureaucrats and that autonomy and performance incentives are complementary.

Additional rationale for rewarding managers is that they can exert substantial influence on overall organizational performance. Through an experiment in Indian textile firms, for instance, Bloom et al. (2011) find that a management intervention raised productivity by 18%. McKenzie

[4] The literature on managerial incentives in the private sector includes, for example, Jensen and Murphy (1990); Hall and Liebman (1998); Hall and Murphy (2003); and Oyer and Shaefer (2005). See Murphy (1999) for a review of CEO pay. For reviews see (Gibbs; Canice Prendergast 1999; R. Gibbons 1998; Murphy 1999; Bloom and Van Reenen 2011).

& Woodruff (2014) review similar management interventions in private firms and paint a more modest picture of these interventions, however.

The quality of school managers (principals), specifically, has long been recognized as an important factor affecting school function anecdotally (Darling-Hammond et al. 2007) and has more recently been empirically tied to student outcomes in developed countries (Branch, Hanushek, and Rivkin 2012). Research has suggested that school managers' management skills, influence on organizational identity, and role in motivating teachers and students all have important implications for school function (Horng, Klasik, and Loeb 2010).

In one recent study, Branch et al. (2012) use value added approaches typically applied to teacher effectiveness studies to investigate the impact of individual school managers on student outcomes. Their estimates suggest a large amount of variation in principal quality and their lower bound (most conservative) estimate suggests that 1 standard deviation increase in manager quality is associated with a 0.05 standard deviation increase in student performance on standardized exams for all students in the school annually. Their results further suggest that one important way through which managers influence school quality is through their influence on existing teachers and on selection of teachers and teacher turnover. This finding has been echoed in other studies on the roles of school management (Boyd et al. 2011).

If, as in developed countries, there is significant variation in school manager quality and this is associated with student outcomes in developing countries as well, this suggests that incentives for school managers in developing country school systems may hold promise as an effective way of improving school system performance and student outcomes.

Chapter 3: Anemia in Rural China: Causes, Consequences, Prevalence and Interventions

The context of this study is school-based programs to reduce iron deficiency anemia in rural Western China. In this chapter, I provide background on the causes and consequences of anemia, the prevalence of anemia in western China and known effective strategies to address iron deficiency.

3.1 Causes and Consequences of Anemia and Iron Deficiency

Anemia is estimated to affect nearly one quarter of all school-aged children (World Health Organization 2001). While there are many causes of anemia (including a variety of genetic disorders, nutritional deficiencies and infections), iron deficiency accounts for around 50% of cases globally (World Health Organization 2001; Balarajan et al. 2011; Pasricha et al. 2013).[1] In China, studies have shown that 85-95% of anemia is due to iron deficiency (Du et al. 2000). Iron deficiency anemia (IDA), however, is a severe form of iron deficiency; iron deficiency also exists without anemia and can be 2 to 3 times more prevalent in a population than IDA (Yip 2001).[2] In China, as in many developing countries, the main cause of iron deficiency is low bioavailability of dietary iron (Du et al. 2000).

The consequences of iron deficiency—with or without anemia—can be substantial, particularly for children during critical developmental periods. A large body of literature links iron deficiency to fatigue and reduced work capacity among adolescents and adults, impaired cogni-

[1] There is some debate in the public health literature on the proportion of the anemia burden attributable to iron deficiency (Balarajan et al. 2011). Intestinal worms are unlikely to be a major cause of anemia in our study areas as the prevalence of hookworm (the parasite most commonly associated with anemia) is low (Xu et al. 1995).

[2] Due to the cost of assessing iron deficiency directly in surveys, anemia prevalence is often used as an indicator for the amount of iron deficiency in a population though it lacks specificity to establish iron status (Balarajan et al. 2011). Anemia is defined based on established cutoffs in hemoglobin concentration, which can be cheaply and reliably assessed in the field (World Health Organization 2001; Balarajan et al. 2011).

tion and cognitive development among children and reduced immune response for all age groups (Thomas et al. 2006; Yip 2001; World Health Organization 2001; Balarajan et al. 2011). More recently, iron deficiency has also been linked to attention deficit hyperactivity disorder (Konofal, Lecendreux, Deron, Marchand, Cortese, Zaim et al., 2008). Likely a result of these effects on cognition and behavior, iron deficient school-aged children have also been shown to have inferior educational outcomes, including grades, attendance and attainment (Taras, 2005; Nokes, van den Bosch & Bundy, 1998). One available estimate of the economic burden of iron deficiency – taking into account lost physical productivity among adults and cognitive effects among children – calculates the median present value of losses in a selected group of 10 developing countries to be $25.60 per capita (Horton and Ross 2003).[3]

3.2 Anemia Prevalence in Western China

This study took place in primary schools across nationally-designated poor counties in three provinces of western China (see Figure 1).[4] Despite rapid economic development and rising incomes in recent years, anemia rates among school-aged children in rural China remain stubbornly high. Approximately one third of children in nationally-designated poverty counties of Northwest China between ages 8 and 12 are anemic (Luo, Wang, Zhang, Liu, Shi, Miller et al., 2011b; Luo, Kleiman-Weiner, Rozelle, Zhang, Liu, Sharbano et al., 2010). [5][6]As in China generally, anemia in this region is largely due to iron deficiency: diets contain insufficient sources of iron (such as meats and green vegetables), as well as vitamin C which promotes iron absorption. [7]

[3]In 2012 dollars, converted from $16.78 in 1994 dollars as reported in the study.

[4]The Chinese government designates a list of "poverty" counties, which it uses to target poverty-alleviation funds. The current list, released in 2012, includes 592 of China's 2,862 county-level divisions. (http://www.globaltimes.cn/NEWS/tabid/99/ID/701105/Chinas-poorest-counties-named.aspx, accessed Sep. 11, 2013).

[5]At altitudes above 1000 meters, it is necessary to adjust hemoglobin concentration when assessing anemia status as the distribution in normal populations increases in response to lower partial pressure of oxygen and reduced blood oxygen saturation (Nestel 2002). To adjust measured hemoglobin, we use the following formula developed by the US CDC: $Hb_{adjusted} = Hb_{measured} + 0.32 * (Altitude(m) * 0.0033) - 0.22 * (Altitude * 0.003)^2$.

[6]Even children with Hb concentration above 120 g/L can be considered at risk of anemia given significant variability in hemoglobin values over time. In particular, there is a significant amount of seasonal variation in hemoglobin measurements, possibly due to seasonal availability of agricultural products.

[7]Helminth infections are unlikely to be a major cause of anemia in our study areas as the prevalence of hookworm (the parasite most commonly associated with anemia) is low (Xu et al. 1995).

Province	Year	Sample Size	Hemoglobin Concentration (g/l)		Anemia Rate
			Mean	Std. Dev.	
Ningxia	2009	4130	125.2	14.5	16.0
Qinghai	2009				40.9
Shaanxi	2008	3661	122.6	11	38.3
Sichuan	2010	453	128.5	12.5	26.6
Guizhou	2010	437			18.7
Shaanxi	2009	2095	127.1	11.3	26.2
Gansu	2010	1630	129.3	12.2	20.6
Ningxia	2010	1016	132.6	12.0	14.1

Table 3.1: Previous Surveys of Anemia Rates in Western China. Data Source: Surveys by the Center for Chinese Agricultural Policy (CCAP) and the Northwest Socioeconomic Development Research Center (NSDRC). Anemic defined as Hb <120 g/L.

3.3 Biomedical Interventions

Iron deficiency can be effectively prevented and treated through interventions using various strategies to deliver iron, which—if effectively implemented—can be extremely cost-effective (Balarajan et al. 2011; Horton & Ross 2003). Iron repletion can improve – and possibly reverse – the detrimental effects of anemia. Improvements in language and motor development have been observed among pre-school age children in East Africa following increased levels of iron (Stoltzfus, Kvalsvig, Chwaya, Montresor, Albonico, Tielsch, et al., 2001). In a meta-analysis of randomized controlled trials that provided iron supplements, Sachdev, Gera & Nestel (2005) find that iron supplements significantly improved the performance of children on tests of cognitive development, especially among children who were initially anemic.

Worldwide, the predominant approach to deliver iron is to fortify dietary staples. Historically, fortifying staples has proven an effective means to address population-level micronutrient deficiencies worldwide (Allen et al. 2006). An attractive feature of this approach is that it largely does not require change in behavior and can be implemented on a large scale. Although staples, such as flour and soy sauce, can be fortified with iron, this may be ineffective in areas (such as Northwest China) where households grow and consume their own food (Allen et al. 2006). Further, in contrast to the success of fortification in addressing other micronutrient deficiencies, such as in iodine and Vitamin A, evidence that fortification can similarly address iron deficiency at the population level is more limited (Uauy, Hertrampf, and Reddy 2002).

An alternative set of strategies focuses on increasing dietary intake of iron. This can be accomplished by encouraging both increased consumption of iron-rich foods and the appropriate mix of other foods that can affect iron absorption. The type of dietary iron most easily absorbed (called heme iron) is found in animal sources (red meats, fish and poultry). Plant food sources (green vegetables, etc) provide a less easily-absorbed type of iron, non-heme iron. The absorption of non-heme iron can be improved, however, by increasing consumption of fruits and vegetables containing vitamin C and reducing consumption of foods that inhibit absorption (such as milk and other calcium-rich products). [8]As noted, the major cause of iron deficiency in China appears to be low dietary iron bioavailability: a significant amount of iron is consumed from plant sources, but this is not easily absorbed and consumption of meats is low, particularly in poor areas of western China (Du et al. 2000). Although encouraging dietary change may be desirable as a long-run solution to iron deficiency, changing dietary behavior through policy interventions is relatively difficult.

Finally, micronutrient supplements (or vitamins) containing iron can be provided at the individual level. Indeed, previous research in China found that providing daily supplements containing iron to children in rural primary schools over five months reduced anemia rates by 10 percentage points (Luo et al. 2012). To be effective, however, regular consumption over time is necessary. Consequently, compliance may be inadequate due to the need for sustained effort. [9]

Although iron deficiency is straightforward and inexpensive to treat in technical terms, approaches that rely on behavior change (i.e., strategies other than centralized fortification of dietary staples—which often do not reach many of the poorest households who grow their own food) face substantial challenges. Individuals and households must choose to consume iron rich foods and/or supplements, and organizations—particularly the public sector—must successfully

[8]Initial stores of iron in the body can also affect absorption of dietary iron (Finch 1994). Because there is an optimal level of iron, the body increases absorption when iron stores are low and decreases absorption when initial stores are high. The returns to a given amount of iron are therefore decreasing in initial iron stores.

[9]Previous trials addressing iron deficiency and anemia have suffered from low levels of compliance or attempted to preempt compliance problems. Bobonis et al. 2006, for example, instructed preschool teachers to provide children with iron therapy for 30 days following health camps but found that only around 18 days were actually administered. The WISE study in Indonesia (Thomas et al 2006) hired facilitators to regularly visit participants and remind them to take their supplements.

deliver them. In the case of schools, administrators must take actions to ensure that their school-based anemia reduction programs are implemented effectively (i.e., taking actions to ensure that children consume supplements or work to improve the diets of children at school and home). Even distributing micronutrient supplements requires effort: water needs to be boiled and cooled, it reduces instruction time and many students do not like taking them and need to be supervised by teachers.

3.4 Addressing Anemia Through School-based Programs

School-based interventions are believed to be among the most cost-effective approaches for delivering health and nutrition services to children in developing countries (Bundy & Guyatt 1996; Jukes et al. 2007; Orazem et al. 2008). Indeed, the 2008 Copenhagen Consensus ranked school nutrition programs among the top ten solutions to address global challenges based on cost-effectiveness.[10] Because developing-country school systems tend to be more developed than public health systems and schools are natural points of contact with school-aged children, school systems provide a platform from which health and nutrition interventions can be delivered at relatively low cost (Bundy & Guyatt 1996; Bundy et al. 2006; Jukes et al. 2007). School-based approaches may be particularly effective in countries such as China where school attendance at the primary school level is nearly universal.

Despite evidence of their effectiveness, however, weak incentives for educators to improve health may keep school-based health and nutrition programs from reaching their full potential. Incentives facing educators in developing countries are often weak in general (cf. World Bank, 2004; Chaudhury, Hammer, Kremer, Muralidaran & Rogers, 2006; Banerjee & Duflo 2006; Duflo & Hanna 2005). Further, even motivated educators may focus on traditional responsibilities over health promotion. Although health promotion and education may be complementary functions, they compete for the attention of finite school resources. Poor educator incentives for improving health may therefore reduce the ability of school-based health interventions to improve student

[10] Copenhagen Consensus 2008 results. http://www.copenhagenconsensus.com/projects/copenhagen-consensus-2008/outcome (Accessed Sep. 11, 2013). Nearly every country in the world seeks to feed at least some of its students through school-based programs (Alderman & Bundy 2011).

health outcomes through reduced compliance or diversion of resources to more traditional functions.

In addition to factors affecting delivery on the part of schools (the supply side), household responses to school-based nutrition programs could also limit their effectiveness. Recent studies have highlighted the potential for education and health inputs provided by schools to crowd-out household investments (Jacoby 2002; Das et al. 2013). However, as discussed in the previous chapter, (outcome-based) supply-side incentives could potentially counteract this effect if incentives encourage more engagement with households with the goals of increasing investments education and nutrition at home.

Chapter 4: Conceptual Framework and Research Questions

In this dissertation, I aim to address three main questions about the use of performance pay for school administrators in the context of school-based nutrition programs. First, I study the impact of administrator incentives tied to student nutrition, rewarding administrators based on school level changes in the rate of anemic students over the school year. Second, I study how the responses of managers to performance incentives vary with budget size. Finally, I examine how performance incentives for school administrators to a.) reduce anemia, c.) improve test scores, or c.) both each affect anemia prevalence and academic performance. In this chapter I discuss conceptual issues related to these three main questions.

4.1 Incentives for School Managers

A newly introduced school-based nutrition program is a prime example of what is, according to Wilson (1989), the most common type of change in government agencies: a new program or task that is added on to existing organizational functions. In the current context, administrators may resist devoting resources to implementing a new nutrition program for reasons beyond mere disutility of effort on the part of administrators. First, implementing a new program may detract organizational effort (of administrators themselves and workers under their direction) away from other school functions. Second, even if additional funds are available, money is fungible and there is an opportunity cost of allocating funds to a nutrition program rather than other uses (particularly when monitoring is difficult).

Managerial incentives (for school administrators) may be particularly salient in this context. Unlike incentives for front line workers (the main purpose of which is to elicit effort in line with the incentivized task), managerial incentives tied to project-specific outputs can also encourage the allocation of resources (funds, effort of subordinates, etc) under their control. Holmstrom and Costa (1986), for example, present a model that highlights the role of managerial decisions regarding project investments and how career concerns of managers can cause misalignment with the objectives of superiors. Thus, providing school administrators with incentives tied to student nutrition may help overcome countervailing incentives and encourage "investment" in the nutrition program.

An important feature of the incentives studied here is that they are based on an output (student anemia prevalence). In contrast to rewarding input use (use of pre-specified health technologies for example), rewarding outputs directly–without specifying the precise inputs to use or strategies to deliver them–strengthens the incentives of managers to use specific knowledge (knowledge of local context and appropriate selections of 'inputs'). Output or outcome-based incentives are therefore, in theory, complementary with managers' control over resources. The degree to which the focus of incentives on outcomes (or on managers vs. front line staff) is beneficial depends on whether this specific knowledge is valuable information in the production of target outcomes. Although outcome based contracts are the norm in education (e.g. teacher incentives tied to student achievement), they tend to be less common in health (Miller & Babiarz 2013).[1]

Another relevant feature of the incentives I study is that they are not only outcome-based, but based on outcomes that are jointly produced (with teachers and other school staff, children, and households). This has two main implications in the context of incentives for school administrators to reduce anemia and raise test scores. First, rewards based on these outcomes will strengthen incentives for administrators to encourage others to contribute to the jointly produced outcome (effectively to bring in extra-marginal resources). Second, an increase in effort on the part of administrators may cause others to reduce theirs (Holmstrom 1982). Such an effect is of particular concern in the context of school-based programs as households could reduce their investment in their children's nutrition or education in response to school-based inputs (Das et al 2011, Jacoby

[1] Leonard (2003) provides examples of traditional healers offering payment contracts contingent on health outcomes in Cameroon.

2002). This could dampen the effects of incentives if administrators believe they will not be rewarded due to offsetting reductions in inputs at home (Leonard 2003). On the other hand, administrators are incentivized to prevent this and may use their specific knowledge to do so.

4.1.1 Why Performance Incentives Might Not "Perform"

In the last decade, several studies the have shown that incentives can be ineffective or even backfire for psychological reasons (see recent surveys by Gneezy et al 2011 and Kamenica 2012). One common argument against performance pay in public sector organizations is that individuals in the public sector tend to be more motivated by internal factors (intrinsic, pro-social and public service motivation). If external incentives crowd out or are substitutes to internal motives, they could be ineffective or have unintended consequences (Langbein 2010; Rebitzer and Taylor 2011; Frey and Jegen 2001; Francois and Vlassopoulos 2008; Benabou and Tirole 2003; Gneezy, Meier, and Rey-Biel 2011). For instance, Benabou and Tirole (2003) note that incentives can be detrimental if they signal that a task is difficult.

Performance incentives may also be ineffective for more mundane reasons such as simply not being strong enough to counter existing institutional incentives. Career concerns, for instance, are thought to exert a comparatively larger influence than direct financial rewards (Gibbons & Murphy 1992). This may be particularly important in the bureaucratic environment that I study (the Chinese school system). 'Bureaucratic incentives' (particularly for managers) are often characterized as being less explicit and instead appealing to the career concerns of individuals as well as a sense of organizational 'mission.' Career prospects tend not to be based on outcomes (such as the success of a project) but on compliance with implicit or explicit rules or constraints (often related to organizational mission—Wilson 1989). This is due to the fact that outcomes are often less observable and less attributable to an individual (Wilson 1989). Thus, administrators may avoid a task if they do not view it as being in line with the primary mission of the organization or believe doing so would violate organizational rules. Bureaucrats themselves also tend to be extremely cautious, either due to self-selection of more risk-averse individuals or the nature of organizational incentives (Wilson 1989; Dixit 2002; Dasgupta & Sarafidis 2009; Buurman et al. 2012). As a result, bureaucrats may avoid a task that they view as uncertain or risky. An unfavorable outcome, in particular, could highlight a deviation from organizational rules or more traditional objectives.[2] The influence of the career concerns of managers may be important in a bureaucratic setting if project outcomes are viewed as uncertain (such as an anemia reduction program), especially given the tendency of bureaucrats to be risk averse.

4.1.2 Incentive Size

A secondary question about managerial incentives tested directly is one of the most basic questions of contract design: How do the effects of P4P incentives vary by incentive strength? Specifically, how do outcomes vary by the "slope" of incentive contracts? Theory suggests that stronger incentives should elicit more effort from agents, certeris paribus. However, empirical evidence from a variety of contexts has shown that small incentives or prices can lead to relatively large changes in behavior (Kremer & Miguel 2007; Thornton 2008; Banerjee et al. 2010; Cohen & Dupas 2010; Karlan et al. 2011; Duflo et al. 2011). The findings in these studies suggest that the existence of incentives or a price may be more important than the actual level of incentives.

On the other hand, other studies have suggested that incentives (small ones in particular) can actually worsen performance because contracts can contain signals and information (for example, difficulty) that may reduce effort (Gneezy and Rustichini 2000; Gneezy, Meier, and Rey-Biel 2011). This would occur if the negative effect of signals contained in incentives are stronger than the price effect (the potential for larger rewards). The relationship between incentive size and effort is even less clear in the case of prosocially-motivated agents. When the incentivized task has a prosocial component, incentives can crowd out or substitute internal motivation. In this case the relative impact of small and large incentives depends in part on how this crowding out effect varies with incentive size. Further, crowding out effects may be larger in the case of rewards based on social outcomes (as opposed to inputs) since these correspond more directly with agents' internal

[2] These characteristics are highly relevant in the Chinese context (Tan 2001). I also found these to be important considerations for administrators in qualitative interviews conducted after the experiment (discussed more fully below).

motivations.

4.2 Resources and Incentives

4.2.1 Resource Effects

There are several reasons to believe that increasing the amount of resources under administrator control alone may be sufficient to meaningfully improve nutritional outcomes absent explicit incentives to do so. First, although funds are fungible, they may have a tendency to "stick" to their intended use (a so-called flypaper effect - Hines and Thaler 1995). Second, administrators may also view the program as sufficiently aligned with organizational mission or professional norms. Third, program uncertainty and connection to career concerns may also be less of a concern for administrators (relative to incentives) if they do not feel others would hold them liable for program outcomes. Fourth, increasing the amount of resources under control of administrators also (in effect) increases their autonomy (or decision making authority - at least over fiscal resources). The public administration literature heavily emphasizes the motivational effects of additional autonomy (Rose-Ackerman 1986; Perry & Wise 1990). Thus, increasing the amount of resources under administrator control may have an effort-inducing effect beyond any effect from an increasing resources alone. Finally, increasing the amount of resources under administrator control may have additional effects on effort due to a sense of obligation. Specifically, being provided with funds (and responsibility) for the specific purpose of reducing anemia may - in itself - induce a sense of obligation to effectively implement an anemia reduction program. Such an effect could be connected to professional norms or organizational mission; however, a more general effect due to being endowed with resources for an expressed purpose could also play a role (to my knowledge this is untested in the literature).

4.2.2 Resource and Incentive Complementarity

A primary interest in this study is how resources (which are increased to fund a new activity) interact with explicit incentives. Most production technologies will exhibit some degree of substitutability between effort and resources. As resources increase school administrators may decrease effort (which is costly in utility terms), if less is needed to achieve the same outcome. Similarly, increasing resources available to administrators could weaken incentives to expend effort to bring in extra-marginal resources (work through households to improve nutrition at home, for instance).

Effort and resources could also be complements in the technology. For example, with enough resources, the administrators could make investments that will increase the marginal productivity of effort. More simply, lack of resources could also merely constrain the feasible set of managerial actions. This is a concept emphasized in the literature: several studies have argued that incentives and control (or autonomy in the organization) are closely linked and that autonomy is a necessary condition for incentives to improve efficiency (Verhoest et al. 2004; Prendergast 2002; Athey and Roberts 2001; Gibbons 2009; Qian 1994; Jensen and Meckling 1976; Aghion and Tirole 1997).

Incentives and resources may also interact for psychological reasons. In the last decade, several studies have shown that incentives can also backfire (see recent surveys by Gneezy et al 2011 and Kamenica 2012). Although this literature analyzes the effect of incentives given a level of resources, some of its insights can also shed light on how incentives and resources may interact. For instance, Benabou and Tirole (2003) note that incentives can be detrimental if they signal that a task is difficult. Presumably this would most likely occur if an agent (a school administrator) believes the principal (education ministry officials) has superior knowledge of the technology producing student nutrition. In the current setting, it is possible that incentives, larger block grants, or their interaction signal task difficulty or complexity. Benabou and Tirole (2006) also note that an individual might not engage with incentivized pro-social tasks. If their peers, friends or neighbors know that an agent has incentives for a task, this creates uncertainty as to the true motivation of the agent's effort. If the agent's original motivation for performing a task is based on image, they may be less likely to provide effort if given an incentive. In our setting a larger grant could increase visibility, amplifying this negative effect of incentives. [3] Note that

[3]Strictly speaking, this effect might also be present if actions are not visible to others, but self-image is important, that is, assessing oneself as in the eyes of an "impartial spectator."

the incentivized task need not be purely pro-social: a similar effect could occur when a task is considered part of an organization's mission.

The general insight of Benabou and Tirole (2006) is that incentives have consequences about how one's actions are interpreted by others as well as themselves. A similar idea is expressed by Heyman and Ariely (2004) who note that monetary incentives can cause individuals to shift from a 'social' to a 'monetary' decision-making frame. Thus a risk averse manager may choose not to invest in a profitable task which has a positive probability of a bad outcome if she believes that incentives (or ability to profit personally) would make the outcome more attributable to her. This may be particularly relevant in bureaucracies where, as discussed, more weight is put on an individuals adherence to organizational norms or rules rather than outcomes (Wilson 1989). Increased resources could exacerbate such an effect since more resources expand the set of feasible risky activities. How incentive reframe a task could also potentially negate or offset any effects that increased resources have on effort. An administrator who otherwise feels an obligation to effectively use resources she is endowed with, for instance, may feel that she can 'pay' to avoid this obligation in foregone incentive pay.

4.3 Multiple Tasks and Incentives

4.3.1 Multitasking and School-based Nutrition Programs

A common feature of public organizations is that they are often charged with multiple functions or roles, often for which success is not easily measured and thus they cannot be contracted upon (Dixit, 2002). In reference to education, Dixit (2002) describes school systems as "multitask, mulitprincipal, multiperiod, near-monopoly organization[s] with vague and poorly observable goals." In such a setting, it is possible that the introduction of performance pay tied to a subset of these functions can refocus resources away from others. School systems are prime examples of organizations with multiple roles. Schools are often charged with: increasing marketable skills of students, instilling behavioral norms, infusing national identity, and promoting health, to name a few. In developing countries, the school's role as a platform for public health interventions is often given added importance as school systems are used to supplement underdeveloped public health systems (Bundy et al., 2006). This implies a multi-tasking problem (a la Holmstrom and Milgrom (1991) and Baker (1992)), suggesting that the relative strength for organizations to devote effort and resources to these roles have important implications for performance. [4]

4.3.2 Incentives and Task Complementarity

I examine multitasking effects of incentives directly in the context of school-based health programs as part of the experiment. I abstract from the issue of observability, however, and focus on the question of how performance incentives based on aligned and misaligned outcomes affect outcomes. Specifically, I examine how anemia-based incentives and test-based incentives each affect anemia and academic performance (exam scores). I also test whether these two incentives are complements or substitutes in the production of student nutrition and academic performance (and the combined effect of the two incentives). [5]

Beyond testing the direct effects of anemia incentives on anemia reduction and test incentives on test scores (and interactions), the experimental design allows me to simultaneously examine how incentives affect more vs less aligned outcomes. Because anemia reduction is an input to academic production, actions taken by principals to reduce anemia should be complementary with improving academic performance. If principals believed that reducing anemia could raise academic performance (as they were told during training), performance incentives tied to test scores may also strengthen incentives to reduce anemia. The converse is not true, however, since academic

[4] Evidence of multitasking effects has been found in a number of empirical studies of service delivery. Rasul and Rogger (2013) show evidence that incentives are negatively correlated with project completion in part because incentives skew the allocation of bureaucratic effort toward more easily observed and contractable projects. Further, this negative correlation is strongest for more complex and 'unfamiliar' projects. Vermeersch and Kremer (2005) find that a subsidized school-based meal program in Kenya crowded out teaching time by 15%. In their study of incentives for local governments to provide health and education services in Indonesia, Olken, Onishi and Wong (2012) find evidence that expenditures were re-allocated from education to health spending; however, they find no decline in education outcomes and attribute this decline in spending to an increase in technical efficiency.

[5] Interest in testing these hypotheses was largely motivated by the findings of the pilot study implemented in 2009-2010 in a smaller number of schools. In that study it was found that 1) anemia reduction incentives had a much larger effect in schools where administrators had pre-existing incentives for good exam scores (Miller et al. 2012) and 2) that incentives for anemia reduction improved test scores for anemic students (presumably through improved health) but had a significant negative effect on scores for students healthy at baseline (Sylvia et al. 2013). These findings suggest that anemia reduction incentives may have crowded out educational inputs (but this was counteracted by improved health for initially anemic students), but that test score incentives and anemia incentives may be complementary in the direction of improved student health. These results however we based on non-experimental variation in existing test score incentives. In this study, I introduce exogenous variation in test score incentives and an arm with both incentives to test complementarity.

performance is not an input to anemia reduction. How test incentives impact anemia, therefore, illustrates the case where incentives are relatively better aligned with a task. Anemia incentives, on the other hand, do not strengthen incentives to raise test scores and any positive effect of anemia incentives on academic performance should solely be due to the effect of improved nutrition.

Consider a basic model of multitasking where school administrators choose (unobservable) effort (including effort and resources), $e \equiv (e^A, e^T)$, which can be devoted to reducing anemia or raising test scores. Assume administrators are faced with a production function for anemia reduction that increases strictly with effort devoted to reducing anemia, $A = f(e^A)$[6][7] and an 'educational production function' that increases with effort devoted to raising test scores and with anemia reduction, $T = g(e^T, A)$. Further, expending effort is costly and this cost is given by $C(e)$ which is strictly convex and continuously differentiable. Increasing effort in one dimension is also assumed to increase the marginal cost of effort in the other, i.e. $C_{TA} > 0$ where $C_{TA} \equiv \frac{\partial C}{\partial e^T \partial e^A}$. Letting P denote the compensation of a school administrator comprised of a base wage and bonuses tied to anemia reduction and test scores,

$$P = \tilde{w} + \beta^A A + \beta^T T$$

The administrator chooses $e^A$and $e^T$to maximize

$$P - C(e) = \tilde{w} + \beta^A A + \beta^T T - C(e)$$

Assume $A = e^A$. Then, the first order conditions are

$$C_A = \beta^A + \beta^T g_A \quad (4.1)$$

$$C^T = \beta^T g_T \quad (4.2)$$

[6]This could also be thought of as a production function for hemoglobin concentration.

[7]Although potential shocks to outcome measures are an important consideration of outcome-based contracts, deterministic production functions suffice for this illustration.

So administrators choose effort in the two dimensions to equate the marginal cost of effort in each dimension with the marginal revenue from increasing that effort. This model has four implications. First, trivially, administrators offered a bonus tied to anemia reduction will increase effort devoted to anemia reduction. Second, adding a bonus for anemia reduction only will lead to a reduction in effort devoted to test scores. This follows from the fact that increasing e^A increases the marginal cost of e^T.[8] Note, however, that the ultimate effect on academic performance is less clear and also depends on how test scores improve with better health, g_A. Third, the effects of a bonus scheme rewarding test scores only is unclear. Because A contributes to the production of test scores (and administrators know this), how increasing β^T affects e^T and e^A not only depends on direct 'price' effects but also on administrators' perception of g_A. How administrators will respond when both dimensions are rewarded is similarly unclear and will additionally depend on cross price effects which are ambiguous (Mullen, Frank & Rosenthal 2010; Sherry, Bauhoff & Mohanan 2013).

[8] To see this, set $\beta^A > 0$ and $\beta^T = 0$. Totally differentiating the first order conditions yields:

$$\begin{aligned} C_{AA} de^A + C_{AT} de^T - d\beta^A &= 0 \\ C_{TA} de^A + C_{TT} de^T &= 0 \end{aligned}$$

the second condition implies that

$$\frac{de^T}{d\beta^A} = -\frac{C_{TA}}{C_{TT}} \frac{de^A}{d\beta^A}$$

which shows clearly that the reduction in effort devoted to test score improvement is proportional to the increase in effort devoted to anemia reduction. Further, in line with common sense, this reduction will increase with administrators' perceptions of how increasing effort devoted to anemia reduction increases the marginal cost of test score effort and will decrease with their perception of how much the marginal cost of test score effort increases with additional units.

Chapter 5: Experimental Design, Data Collection and Empirical Approach

5.1 Experimental Design

5.1.1 Power Calculations

Power calculations were conducted before the beginning of the trial by Montecarlo simulation using parameter estimates from a 2010 pilot study. From the pilot study data, the intraclass correlation was estimated using a subsample of anemic students and adjusting for a number of covariates (lag value of the outcome variable, number of students in schools, whether the school has a canteen, student-teacher ratio, distance to the furthest village served by the school, percent boarding students, county dummies). After adjustment, the intraclass correlation at the school level for hemoglobin was found to be 0.11.

In addition, the following parameters were assumed:

- 15% loss to follow-up
- 50 children tested per school
- 13 anemic at baseline tested at follow-up accounting for loss
- 30.6% anemia rate (at 120 g/L cutoff)

Monte Carlo simulations using 500 replications were run on the regression:

$$Hb_{is} = \alpha + \beta_1 Test_s + \beta_2 Anemia_s + \beta_3 (Test_s) \times (Anemia_s) + \varepsilon_{is} \quad (5.1)$$

using the subsample of anemic students at baseline in the Subsidy and Anemia Incentive arms of the pilot study. Where $Test_s$ is an indicator for whether the school administrator had a (non-experimental) performance incentive tied to student test scores and $Anemia_s$ is an indicator for whether the school was in the Anemia Incentive experimental group.

Based on these simulations, 260 schools with 65 schools in each arm give 96% power to detect a test incentive effect (β_1) of 0.25*sd, 100% power to detect an anemia incentive effect (β_2) of 0.35*sd, and 57% power to detect a coefficient on the interaction term (β_3) of -0.2*sd (all at 5% significance) where sd = the standard deviation of the residual = 1.03762. Note that this estimate did not take into account gains from stratification.

Power calculations were also done for test score outcomes. However, given the higher ICC, more power was needed for anemia outcomes so these were used to determine sample size.

5.1.2 Sampling

A canvas survey was conducted in August 2011 in 36 nationally-designated poverty counties in five regions of Qinghai, Gansu, and Shaanxi provinces (Shown in Figure 5.1). This survey yielded a list of all primary schools and the number of students enrolled in each as reported by the local education bureaus. From this list, schools were eliminated that had less than 150 or more than 300 students.[1] The remaining list served as the sampling frame, and 300 of these schools were randomly selected for inclusion in the study. To minimize the possibility of contamination, only one school was selected per township. [2]

5.1.3 Allocation

Experimental allocation was done using a 3-by-2 factorial design. After conducting our baseline survey in September 2011 (at the beginning of the 2011-2012 academic year), I randomly assigned each of 300 study schools to one of the ten experimental cells shown in Figure 5.2. The five paths show the randomly-assigned incentive groups: a group without incentives (Group A),

[1] A lower bound of 150 students was chosen to ensure that the number of samples students per school was enough to meet power requirements. 300 was chosen as the upper bound to keep the project within budget. These bounds are on reported school sizes; actual numbers of students are often significantly less than reported. Note that 39.9% of rural primary schools in the sampling frame (all rural primary schools in project counties) were reported to be within this range.

[2] Local administration of schools is generally done at the school district level, which is below the township. Contamination due to two school administrators meeting at events organized at higher levels, for example, was thus unlikely.

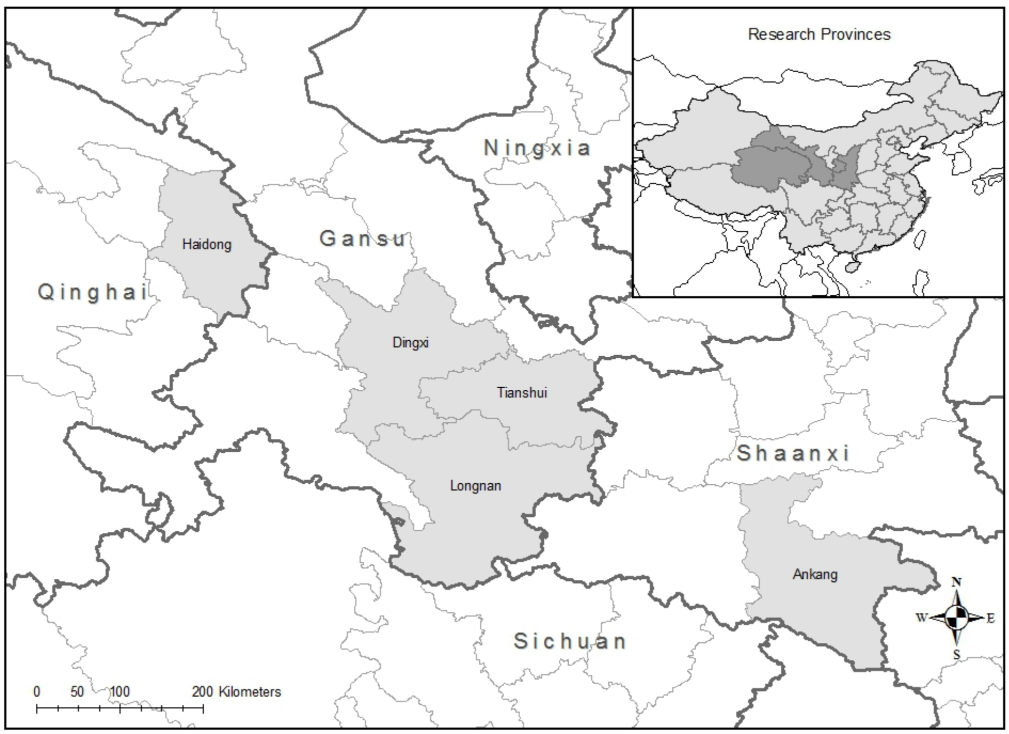


Figure 5.1: Study Regions

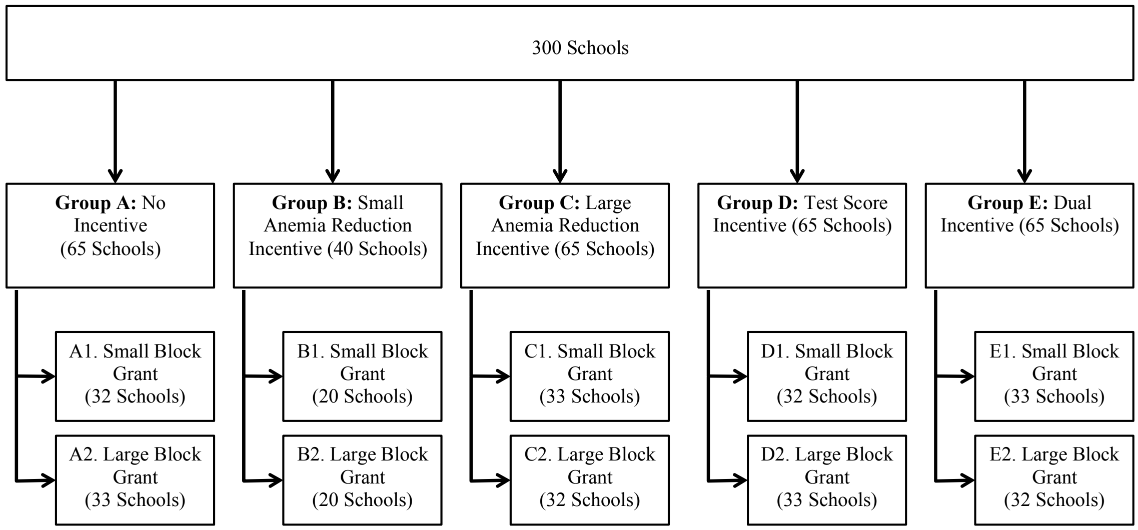


Figure 5.2: Experimental Design

a "small" anemia incentive group (Group B), a "large" anemia incentive group (Group C), a test score incentive group (Group D), and a dual incentive group (Group E). I orthogonally crossed these three incentive arms with two block grant groups: a "small" block grant group (Group 1) and a "large" block grant group (Group 2). Because knowledge about anemia in our study areas was negligible, we conducted a health education campaign about anemia in all study cells (described in detail in below).[3] The primary reference group in this ten-cell design is the default policy (education about anemia coupled with a modest budget transfer and no incentives), Group A1. [4]

To improve power, I used a stratified randomization procedure. Specifically, the randomization process was stratified on baseline mean school-level hemoglobin concentration and mean combined standardized math and Chinese exam scores. Each of these variables was divided into 5 strata yielding 25 total blocks. The analysis takes this randomization procedure into account (Bruhn & McKenzie 2009).

A map of study sites is shown in Figure 5.3.

[3]In the pilot study, we found that educating school administrators on anemia (providing the same information as in the current study) alone, without incentives or grants, had no detectable impact on anemia rates.

[4]This is, for example, how a recent school feeding policy, to which the government has appropriated 16 billion yuan annually, has been implemented in many areas in practice. Under the program, local education bureaus and schools receive 3 yuan per day per student (4 yuan for boarding students) to provide nutritious meals. How exactly the program is implemented and monitored varies widely across localities.

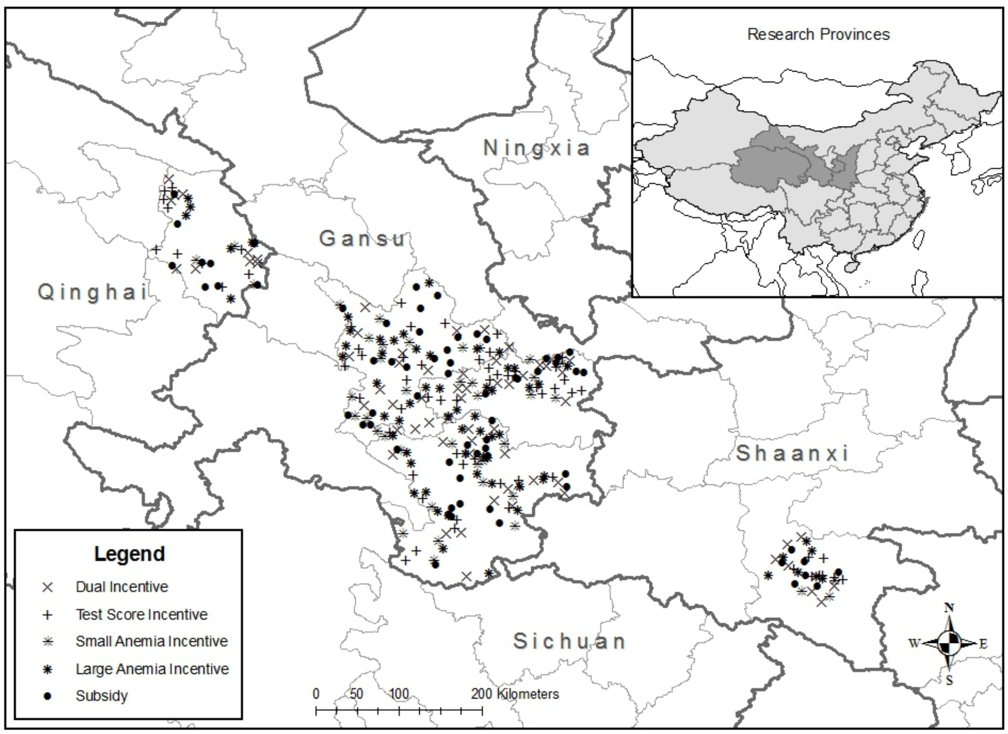


Figure 5.3: Study Sites

5.1.4 Interventions

5.1.4.1 Health Training on Anemia

Because knowledge about anemia in the study areas was negligible, all 300 school administrators in the study were given training on iron deficiency anemia prior to revealing their randomly-assigned study cell. Training materials were based on published, peer-reviewed studies and specifically included information about: a) the prevalence and causes of anemia, b) the consequences of anemia (including its effect on cognitive development and academic performance), and c) efficacious nutritional approaches to reduce anemia (increasing dietary intake of iron-rich foods, nutritional supplementation with iron fortified soy and flour or with supplements, etc.). This training was given through a presentation by a nutrition expert from a local university [5] and reference materials given to school administrators.

5.1.4.2 Block Grants

Across the five incentive arms, schools were randomly assigned to one of two block grant groups: small (Group 1 in Figure 5.2) or large (Group 2 in Figure 5.2). Although funds were given in the context of the nutrition program roll out, administrators were free to allocate these to other school functions at their discretion. The small grant was 0.3 yuan ($0.05)[6] per student per day (85 schools), which is adequate to purchase a supplement for each student to take daily. The large grant was 0.7 yuan per student per day ($0.11) (85 schools). In total, small block grant schools received 7,452 yuan ($1,183) on average and large block grant schools received 17,388 yuan ($2,760). For reference, school budgets in the project area were approximately 500 yuan per student per year (or 4 yuan per day).

These grants were given to schools in two installments, once at the beginning of the program and another approximately half way (see Figure 5.4). Because a market for iron supplements and multivitamins did not exist in many project areas, administrators were given the option to purchase

[5] A translation of the transcript is in the appendix.

[6] I use a conversion rate of $1 = 6.3 RMB throughout.

vitamins from the Chinese Center for Agricultural Policy at cost. They initially decided how much to purchase when they first learned of the grant and filled out their "planning budget" (see data collection below). Administrators were aware that they would receive vitamins in kind and the remainder of the grants in cash. Prior to the distribution of the first installment of the grant, administrators were contacted to confirm their vitamin order. They were also told that they could make contact anytime to change vitamin orders if they chose to do so. The same procedure was followed for the second installment at the beginning of semester 2.

5.1.4.3 Incentives

The incentive arms were as follows. In the *small anemia-reduction* incentive group (40 schools, Group B in Figure 5.2), administrators were given contracts detailing a financial incentive to be paid to them as private income according to the net reduction in number of sample students with anemia between the beginning and end of the school year. The specific structure of the small incentive contract was:

$$Payout_{SI} = 12.5yuan \times (N_b - N_e) \; if \; N_b - N_e > 0$$

where N_b is the number of sample students found to be anemic at baseline and N_e is the number of these same students who were anemic at the time of the endline survey. The contract increment (12.5 yuan (RMB) per student reduction– approx. $2) was set using experience from the pilot study to provide an additional 0.2 months of a principal's annual salary for a feasible reduction in anemic students.[7] The small size of this incentive was chosen so that rewards themselves were of little monetary value.

The *large anemia reduction* incentive group (65 schools, Group C in Figure 5.2) was identical to the small incentive arm except that the magnitude of the incremental incentive was ten times greater (125 RMB, or about $20):

[7] There are presumably superior contract structures, however this structure was chosen for simplicity. Readily-understandable contracts may appear more transparent to school principals and promote credibility.

$$Payout_{LI} = 125yuan \times (N_b - N_e) \; if \; N_b - N_e > 0$$

This contract increment implies 2 additional months of annual salary for a feasible reduction in anemia prevalence. If the anemia rate increased over the school year, administrators did not receive a payout.

Administrators in the *test incentive* group (65 schools, Group D in Figure 5.2) were given an incentive contract stipulating rewards based on the average gain in student scores on standardized exams over the course of the school year as follows:

$$Payout_{TI} = 800yuan \times (\bar{S}_e - \bar{S}_b) \; if \; \bar{S}_e - \bar{S}_b > 0$$

For every 1 point gain in the average of student scores on standardized exams in math and Chinese (combined) principals received 800 yuan (approx. $125). This was calibrated such that a reward for a 0.2 SD increase was approximately equal to one month's salary. Principals were not penalized for a reduction in average exams scores.

The final incentive group, the *dual incentive* group (65 schools, Group E in figure 5.2), received both the *large anemia reduction* and *test incentives.* An example of this dual contract is in the appendix.

5.1.5 Implementation

At the time of treatment, principals first received training on anemia. Following this, research team members revealed randomly-assigned study groups to principals, told them how many sampled students in their schools were anemic at baseline (the identity of these children was not revealed), and signed incentive contracts with principals. Contracts were written on Chinese Academy of Sciences letterhead (which is part of the Chinese Government) and counter-signed by the deputy director of the implementing research center (principals signed two copies of the contract, one of which they took with them). Contracts were also stamped with the official seal of the Chinese Center for Agricultural Policy. At the same time, administrators were informed of the grant they would receive and made an initial decision on what portion they would spend on multivitamins. Note that all interventions were implemented in partnership with the local education bureaus. Beyond being a necessity in China, doing so also creates the impression to participants that interventions were a government program.

5.2 Data Collection

5.2.1 Sampling

Data collection efforts within schools focused on 4th and 5th grade students.[8] Within each school all 4th and 5th grade students were given student surveys (discussed below). For hemoglobin tests, 50 fourth and fifth grade students were randomly selected.[9] To minimize the possibility that administrators with performance incentives later reallocated resources based on our baseline sample, all students in the school were administered surveys in the baseline. Students in other grades were also sampled for physical examinations. Hemoglobin tests were given to children in sixth grade, but only height and weight was collected for children in grades 1-3.

5.2.2 Surveys

We conducted our baseline survey in September 2011 (at the beginning of the 2011-2012 academic year), implemented our experimental interventions in October 2012, and conducted our follow-up survey in May 2012 (at the end of the same academic year).

Both rounds collected detailed information on students, schools, school administrators, teachers, and households. Student surveys included modules on student background, various health inputs received from schools possibly related to anemia reduction activities (meals eaten at school, distribution of supplements, etc), academic activities including questions on teacher behavior, and

[8] 4th and 5th grade students are old enough to fill out tests and survey instruments yet too young for a significant number to have gone through puberty, which can complicate interpretation of hemoglobin readings.

[9] When there were less than 50 students in the two grades, all students were tested.

general health. To collect information on school and home feeding practices, students were also given food frequency questionnaires asking about food consumption at school and at home over the past week. Although information on student background (including household characteristics) were included in student forms when possible, a household survey was also administered to collect additional information on household behavior (including anemia-related and academic investments) and household characteristics that were unlikely to be known by students themselves.[10] School surveys collected basic information on schools, school finance, and kitchen operations. The teacher survey mainly collected information on teacher background and teaching practices. Homeroom teachers (banzhuren) were given an additional module on student health and school-based health and nutrition related activities.

The school administrator survey collected detailed background information on the school administrator as well as other types of information pertinent to incentive contracts. The basic background information on school administrators included education, teaching experience, administration experience, current structure of pay. In addition, school administrator surveys included a module designed to attempt to measure different subjective perceptions of school administrators that could influence responses to performance pay (including risk preferences, discounting, and social preferences (intrinsic and pro-social motivation). [11]

Several additional types of information were also collected. First, immediately after principals signed incentive contracts and learned of block grants, they were asked to fill out a non-binding budget were they chose how to allocate program funds across different items. Budget items included 20 different nutrition-related (food, supplements, etc.) and education-related (teacher training, books, computers, etc) inputs. Administrators could also write in whatever they wished to purchase. Note that it was emphasized that this budget was non-binding and administrators could

[10]For funding reasons, household forms were not administered by enumerators. These forms were sent home with children to fill out with their parents. Once completed these were returned to homeroom teachers who were provided with pre-paid envelopes to mail these to the implementing research center.

[11]We measured intrinsic motivation (motivation from enjoyment of the act of doing one's job) and pro-social motivation (motivation from desire to affect a pro-social outcome) using psychological scales developed in Grant (2008) translated into Chinese. Although we have not formally established the validity of these scales in China, the questions are not particularly culturally or linguistically sensitive and the measures co-vary with other characteristics in our data in expected ways.

change their mind at any time. [12] An example planning budget is shown in the appendix.

Data on school system characteristics was collected using a survey of school district superintendents. This survey was conducted by telephone at the time of the baseline survey. School district superintendents were asked questions regarding school district characteristics and school district policies regarding the management of school administrators.

5.2.3 Primary Outcomes

The primary outcomes for the trial are student hemoglobin concentration (Hb) (and anemia defined as altitude-adjusted Hb<120 g/L) and student performance on standardized exams in math and Chinese. To collect hemoglobin concentration measurements, nurses from Xi'an Jiaotong Medical School accompanied enumerators during the baseline and endline surveys. Hemoglobin levels were measured on-site (at schools) using HemoCue Hb 201+ systems – a procedure considered state-of-the-art.[13] Before surveys, nurses were extensively trained on the appropriate method of drawing and testing capillary blood samples.

To assess gains in student achievement, standardized exams in math and Chinese were developed in collaboration with local education bureaus based on the national curriculum. One-half of the students in each class were randomly assigned to take either the math or Chinese exam. Questions for the math exam were taken from the question bank of the Trends in International Mathematics and Science Study (TIMSS) and the Chinese test was designed with the help of the Shaanxi Provincial Education Bureau. The exams were extensively pretested to ensure they adequately captured variation in student achievement. In analysis exam scores are normalized by the control group distribution.

[12] This budget was also used to coordinate orders for iron supplements to be delivered to schools. Administrators were free to change their supplement orders at any time before vitamins were delivered.

[13] Each student was only tested once during each round of hemoglobin testing. Although taking multiple measurements per round can help address classical measurement error present in finger-prick assessments, pretesting suggested that efficiency gains from adding another test per subject did not justify the extra cost.

5.2.4 Semi-structured interviews

In addition to quantitative survey and test data, qualitative interviews were also conducted in a convenience sample of 12 schools following the endline survey. Although these schools were a convenience sample, schools were evenly distributed across trial arms and across the two largest sample prefectures (both in Gansu province). Using semi-structured interviews, administrators were asked about their work generally, existing motivations and explicit incentives, how they chose to implement the nutrition program and their impression of the program generally.

A timeline for data collection and intervention activities is shown in Figure 5.4.

5.3 Empirical Approach

Given random assignment schools to treatment cells shown in Figure 5.2, comparisons of outcome variable means across treatment groups provides unbiased estimates of impact of each experimental treatment. However, to increase power, I control for covariates and estimate treatment effects using OLS. These covariates are the same used to conduct power calculations using data from a pilot project. In addition, all of the analyses presented (including outcome variables, regression specifications and hypotheses to be tested) were pre-specified in an analysis plan written and filed before endline data were available for analysis. [14]

As specified in the pre-analysis plan, I estimate the effect of group assignment on child-level outcomes using the following ANCOVA specification:[15]

$$Y_{isct} = \alpha + T'_{sc}\beta + \theta Y_{isc(t-1)} + X'_{isc(t-1)}\gamma + \mu_c + s_{sc} + \varepsilon_{isct} \tag{5.2}$$

where Y_{isct} is the outcome (hemoglobin concentration in g/L, anemia status, or standardized exam score) for child i in school s in county c measured at endline; T_{sc} is a vector of treatment

[14] This analysis plan was filed with the Abdul Latif Jameel Poverty Action Lab at MIT. In instances that deviate from the pre-specified plan due to data constraints or lack of foresight, I state so explicitly in the text and tables.

[15] McKenzie (2011) discusses gains in efficiency from using ANCOVA over a difference-in-difference estimator in estimating treatment effects from randomized experiments. In particular, the ratio of the DID variance to the ANCOVA variance is $2/(1+\rho)$whereρ is the autocorrelation of the outcome variable. Autocorrelations for test scores are typically about 0.5 to 0.6 meaning that we would need a sample size 25-33% larger to achieve the same power using DID.

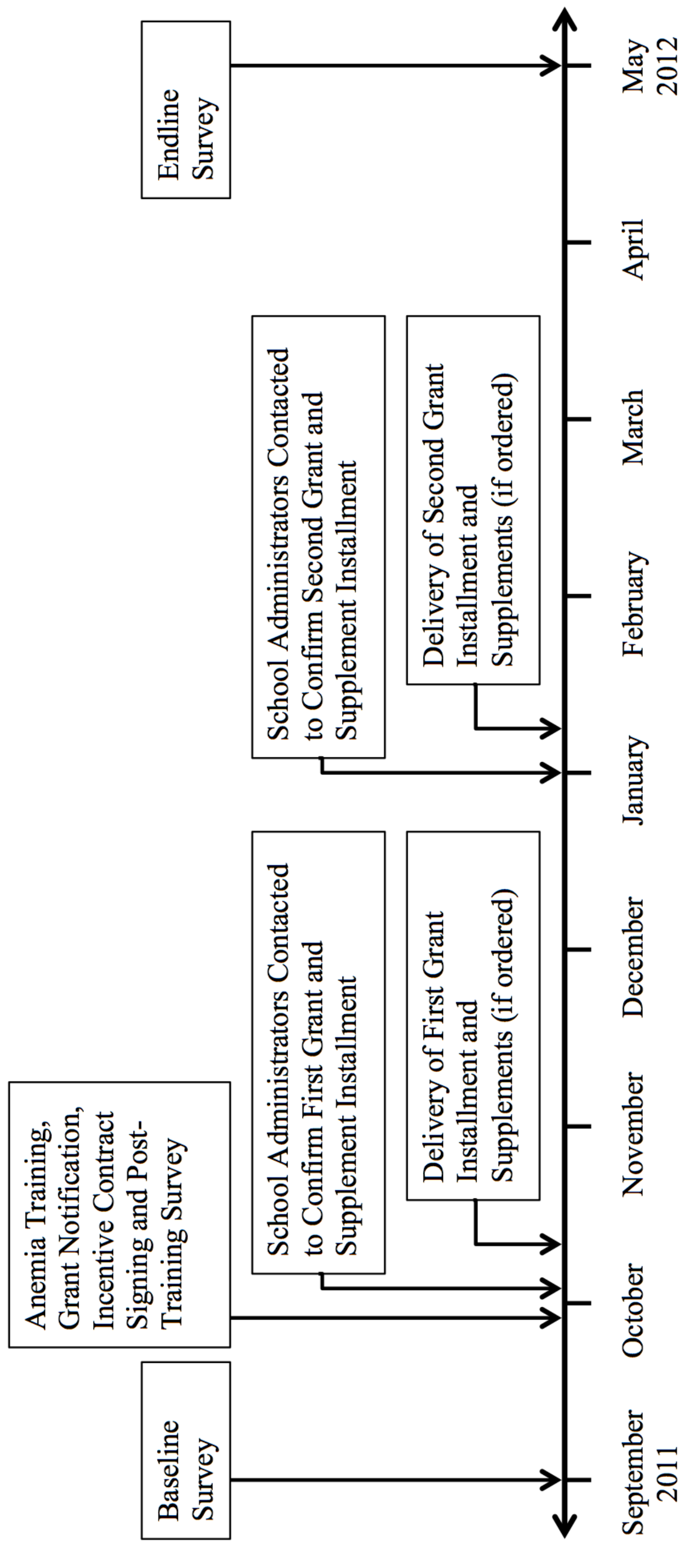


Figure 5.4: Timeline

group dummies; $Y_{isc(t-1)}$is the baseline value of the dependent variable; $X_{isc(t-1)}$ is a vector of child controls (age, class-year, and gender) and school controls (number of students, student-teacher ratio, whether the school has a kitchen, proportion of boarding students, and distance to the furthest village in the school's catchment area); $\mu_c$are county fixed effects; and $s_{sc}$are dummy variables for randomization strata. Standard errors are adjusted for clustering at the school level using the cluster-corrected Huber-White estimator.

For the sake of presentation, I split the analysis to assess specific research questions. To do this, I use a subset of treatment dummies and limit the sample to the appropriate treatment groups and the no incentive group. To estimate the treatment effects of large and small anemia reduction incentives and how they interact with grant size, I use the following specification:

$$Y_{isct} = \alpha + \beta_1 SI_s + \beta_2 LI_s + \beta_3 LG_s + \beta_4 SI_s \times LG_s + \beta_5 LI_s \times LG_s + Z'_{isc(t-1)}\zeta + \varepsilon_{isct} \qquad (5.3)$$

SI_s is a dummy that equals 1 if the administrator in school j was assigned to receive a small incentive contract and 0 otherwise; LI_s is equal to 1 if the administrator in school j was assigned to receive a large anemia reduction incentive contract; LG_s is equal to 1 if the school received a large block grant. $Z_{isc(t-1)}$includes all controls in Equation (5.2). Here, the estimation sample is limited to the comparison (no incentive) group and the anemia incentive groups (Groups A,B and C in Figure 5.2). This specification compares the default policy (i.e., the group that received training, a small block grant and no incentive) to each of the other experimental groups: β_1 gives the additional impact of adding a small incentive to a small block grant and information (Group B1 vs A1 in Figure 5.2); β_2 gives the impact of adding the large incentive (Group C1 vs A1 in Figure 5.2); β_3 gives the impact of additional resources with the large block grant (Groups A2 vs. Group A1 in Figure 5.2). Also of interest are coefficients β_4 and β_5 on the interactions between the large grant and small and large incentives. These coefficients give the additional effect of incentives in the presence of the large block grant compared to their effect when combined with the smaller block grant.

To estimate the effects of test incentives and dual incentives and compare these to anemia incentives, I estimate

$$Y_{isct} = \alpha + \beta_1 AI_s + \beta_2 TI_s + \beta_3 (AI_s) \times (TI_s) + Z'_{isc(t-1)}\zeta + \varepsilon_{isct} \tag{5.4}$$

where $AI_s$is a dummy for all schools with a (large) incentive for anemia reduction and TI_s is a dummy for schools where administrators were given incentives tied to student exam performance. Note that AI_s and TI_s are both 1 for schools in the "dual incentive" group (Group E in figure 5.2). I estimate this equation using all schools in the trial excluding only the "small" anemia reduction incentive group (Group B). I estimate this for the combined sample including both the large and small grant groups (Groups 1 and 2) and also for each group separately. $\beta_1$estimates the impact of anemia reduction incentives, $\beta_2$estimates the impact of incentives for test scores, and $\beta_3$tests whether the two incentives are complements or substitutes. In addition to the individual estimates for $\beta_1 - \beta_3$, I also present pairwise Wald tests of $H_0 : \beta_1 = \beta_2$ and the combined effect of the incentives, $\beta_1 + \beta_2 + \beta_3$, corresponding to the effect of the dual incentive group.

In addition to the primary outcomes (hemoglobin concentration and anemia), I use the same specifications (adjusted appropriately for school-level variables) to estimate effects on secondary outcomes in order to assess the mechanisms through which incentives and block grants worked. All results presented for secondary outcomes were specified in the per-analysis plan. In the pre-analysis plan, secondary outcomes were grouped corresponding to specific hypotheses. In addition to estimating impacts for each variable separately, I also estimate effects on a summary index for each hypothesis group. These indices were constructed using the GLS weighting procedure described in Anderson (2008). For each individual, a new variable, $\bar{s_{ij}}$is created that is the weighted average of the k outcome variables y_{ijk} in hypothesis group, j. The weight place on each outcome variable is set equal to the sum of its row entries in the inverted covariance matrix for hypothesis group j such that

$$\bar{s_{ij}} = \left(\mathbf{1}'\hat{\sum_{\mathbf{j}}}^{-1}\mathbf{1}\right)^{-1}\left(\mathbf{1}'\hat{\sum_{\mathbf{j}}}^{-1}\mathbf{y}_{ij}\right)$$

where $\mathbf{1}$is a column vector of 1s,$\hat{\sum_{\mathbf{j}}}^{-1}$ is the inverted covariance matrix, and $\mathbf{y}_{ij}$ is a column vector of all outcomes for individual i in hypothesis group j. This procedure can have the added advantage of increased efficiency since outcomes that are highly correlated receive less weight and outcomes less correlated (containing new information) receive more weight. Further, $\bar{s_{ij}}$can still be created for observations with a subset of missing outcomes, these outcomes merely receive less weight in the index.

5.4 Summary Statistics, Balance and Attrition

5.4.1 Descriptive Statistics and Balance

Summary statistics and tests for balance across study arms are shown in Tables 5.1-5.4. Tables 5.1 and 5.2 show summary statics and balance tests for the anemia incentive groups and cross-cutting block grant groups for the full sample and sample of children anemic at baseline, respectively. Tables 5.3 and 5.4 show the same for the large anemia incentive, test and dual incentive groups for schools that received small block grants and large block grants separately. In each table, Panel A shows student level characteristics, Panel B shows characteristics of schools, and Panel C shows characteristics of school administrators. The first two columns of the table give the mean and standard deviation of each variable in the comparison group. Columns (3) – (7) show coefficients on treatment variables and interactions estimated using equation (5.1), controlling only for randomization strata fixed effects. The final column shows the p-value from a test that the coefficients are jointly zero.

The number of statistically significant tests is not greater than would be expected by random chance.

	Small Grant, No Incentive		Coefficient (standard error) on:						
	Mean	SD	Small Incentive	Large Incentive	Large Grant	(Small Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	N	P-value: Equality of All Groups
A: Child Characteristics									
1. Hemoglobin Concentration (g/L)	134.191	12.912	-0.912 (1.127)	-1.192 (1.009)	0.514 (1.028)	0.140 (1.501)	-0.021 (1.476)	8398	0.541
2. Anemic (0/1)	0.233	0.423	0.024 (0.017)	0.017 (0.019)	-0.015 (0.018)	-0.001 (0.024)	0.003 (0.025)	8398	0.222
3. Age (years)	10.713	1.173	-0.172 (0.128)	-0.041 (0.111)	-0.030 (0.106)	0.352* (0.185)	-0.013 (0.144)	8398	0.379
4. 5th Grade (0/1)	0.531	0.499	-0.002 (0.006)	0.001 (0.006)	-0.005 (0.008)	0.007 (0.011)	0.001 (0.010)	8398	0.941
5. Female (0/1)	0.485	0.500	0.003 (0.020)	-0.008 (0.017)	-0.009 (0.019)	0.024 (0.030)	0.010 (0.025)	8398	0.808
B: School Characteristics									
6. Number of Students	207.094	64.823	-1.276 (17.567)	3.623 (14.959)	-5.396 (16.043)	25.344 (25.554)	12.357 (20.856)	170	0.797
7. Has Kitchen (0/1)	0.063	0.246	0.141 (0.101)	0.074 (0.075)	0.059 (0.083)	-0.075 (0.162)	-0.068 (0.120)	170	0.681
8. Student-Teacher Ratio	16.228	4.227	2.538* (1.354)	0.893 (1.210)	-0.286 (1.159)	-1.506 (1.911)	1.064 (1.657)	170	0.257
9. Time to Furthest Village Served (mins)	62.031	36.695	12.218 (13.109)	-2.281 (11.564)	3.878 (12.945)	-7.346 (21.467)	3.764 (17.794)	170	0.921
10. Percent Boarding Students (%)	5.327	11.404	1.511 (4.112)	0.106 (3.006)	0.610 (3.492)	-0.079 (6.293)	-1.611 (5.179)	170	0.991
C. School Administrator Characteristics									
11. Male (0/1)	0.938	0.246	0.015 (0.058)	0.056 (0.041)	0.065 (0.042)	-0.012 (0.059)	-0.093* (0.051)	170	0.488
12. Age (years)	39.313	7.253	1.883 (2.047)	1.620 (1.777)	1.892 (1.831)	-5.022* (2.957)	-0.399 (2.560)	170	0.351
13. Higher Education Degree (0/1)	0.906	0.296	0.002 (0.089)	-0.022 (0.078)	-0.122 (0.090)	0.047 (0.133)	0.010 (0.122)	170	0.506
14. Administration Experience (years)	8.031	6.141	-0.242 (1.472)	1.088 (1.729)	0.838 (1.569)	-2.706 (2.156)	-0.310 (2.520)	170	0.141

NOTES. Data source: authors' survey. Table uses full baseline sample. First two columns show the mean and standard deviation of the baseline characteristic in the comparison (small grant, no incentives) cell. Hemoglobin concentration and anemia statistics are shown for the sample of children randomly selected for hemoglobin tests. The remaining characteristics in Panel A are shown for all children in the full sample regardless of whether they were selected for the hemoglobin test. A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). Columns 3 through 7 show coefficients and standard errors (in parentheses) from a regression of the characteristic on incentive and block grant treatment group dummy variables and there interaction estimated using equation (1) but controlling only for randomization strata fixed effects. The final column shows the p-value from a Wald test that coefficients are jointly zero. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 5.1: Descriptive Statistics and Balance: Anemia Incentives (Full Sample)

	Small Grant, Small Incentive		Coefficient (standard error) on:						
	Mean	SD	Small Incentive	Large Incentive	Large Grant	(Small-Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	N	P-value: Equality of All Groups
A: Child Characteristics									
Hemoglobin Concentration (g/L)	118.446	7.541	-0.965 (1.326)	-1.525 (1.163)	-0.653 (1.438)	2.868 (1.959)	1.479 (1.761)	2051	0.420
Age (years)	10.514	1.153	0.046 (0.166)	0.077 (0.125)	0.113 (0.134)	-0.002 (0.242)	-0.070 (0.189)	2051	0.914
5th Grade (0/1)	0.468	0.500	0.055* (0.032)	-0.002 (0.029)	-0.003 (0.031)	-0.106** (0.050)	0.016 (0.042)	2051	0.177
Female (0/1)	0.530	0.500	0.003 (0.044)	-0.021 (0.035)	-0.009 (0.039)	0.001 (0.060)	0.044 (0.052)	2051	0.945
B: School Characteristics									
Number of Students	203.733	55.788	2.424 (16.959)	7.060 (14.194)	-1.925 (15.304)	21.948 (25.245)	9.631 (20.780)	167	0.725
Has Kitchen (0/1)	0.067	0.254	0.135 (0.099)	0.068 (0.077)	0.054 (0.085)	-0.071 (0.161)	-0.052 (0.120)	167	0.732
Student-Teacher Ratio	16.192	4.356	2.859** (1.377)	1.190 (1.216)	0.019 (1.182)	-1.804 (1.928)	0.866 (1.678)	167	0.185
Time to Furthest Village Served (mins)	61.167	37.570	12.294 (13.474)	-2.256 (11.962)	4.020 (12.520)	-7.468 (21.139)	4.605 (17.681)	167	0.918
Percent Boarding Students (%)	4.277	9.493	2.228 (3.976)	0.756 (2.899)	1.310 (3.400)	-0.757 (6.204)	-1.804 (5.107)	167	0.985
C. School Administrator Characteristics									
Male (0/1)	0.967	0.183	-0.015 (0.051)	0.028 (0.032)	0.038 (0.034)	0.014 (0.053)	-0.070 (0.046)	167	0.606
Age (years)	39.567	7.398	1.550 (2.112)	1.299 (1.837)	1.599 (1.882)	-4.730 (3.022)	0.090 (2.601)	167	0.383
Higher Education Degree (0/1)	0.900	0.305	0.018 (0.092)	-0.007 (0.081)	-0.107 (0.093)	0.032 (0.136)	-0.007 (0.126)	167	0.558
Experience (years)	8.333	6.227	-0.194 (1.531)	1.124 (1.786)	0.898 (1.630)	-2.761 (2.210)	-0.165 (2.577)	167	0.137

NOTES. Data source: authors' survey. Table uses sample of children anemic at baseline. First two columns show the mean and standard deviation of the baseline characteristic in the comparison (small grant, no incentives) cell. Hemoglobin concentration and anemia statistics are shown for the sample of children randomly selected for hemoglobin tests. The remaining characteristics in Panel A are shown for all children in the full sample regardless of whether they were selected for the hemoglobin test. A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). Columns 3 through 7 show coefficients and standard errors (in parentheses) from a regression of the characteristic on incentive and block grant treatment group dummy variables and there interaction estimated using equation (1) but controlling only for randomization strata fixed effects. The final column shows the p-value from a Wald test that coefficients are jointly zero*, **, and *** indicate significance at 10%, 5% and 1%.

Table 5.2: Descriptive Statistics and Balance: Anemia Incentives (Anemic at Baseline Sample)

	Control		Coefficient (standard error) on:			N	P-value: Equality of All Groups
	Mean	SD	Anemia	Test	Dual		
A: Child Characteristics							
1. Hemoglobin Concentration (g/L)	133.955	12.806	-1.053 (0.725)	-0.648 (0.743)	0.216 (0.666)	12878	0.282
2. Anemic (0/1)	0.232	0.422	0.019 (0.012)	0.010 (0.011)	0.009 (0.011)	12878	0.503
3. Baseline Exam Score (Normalized)	-0.000	1.000	0.004 (0.021)	-0.005 (0.021)	-0.024 (0.021)	18118	0.549
4. Age (years)	10.691	1.167	-0.035 (0.073)	-0.033 (0.077)	-0.047 (0.070)	18124	0.925
5. 5th Grade (0/1)	0.527	0.499	-0.008 (0.010)	-0.013 (0.010)	0.005 (0.011)	18124	0.359
6. Female (0/1)	0.485	0.500	0.002 (0.011)	-0.002 (0.011)	0.005 (0.012)	18123	0.961
B: School Characteristics							
7. Number of Students	204.354	61.355	10.442 (10.323)	0.397 (9.777)	7.175 (10.521)	260	0.643
8. Has Kitchen (0/1)	0.108	0.312	0.046 (0.059)	0.043 (0.061)	0.065 (0.061)	260	0.739
9. Student-Teacher Ratio	15.906	4.627	1.497* (0.808)	0.684 (0.833)	1.911** (0.849)	260	0.112
10. Time to Furthest Village Served (mins)	64.385	47.331	-1.214 (8.415)	0.828 (8.722)	8.672 (9.217)	260	0.701
11. Percent Boarding Students (%)	5.913	12.974	-0.072 (2.384)	0.606 (2.804)	0.935 (2.634)	260	0.975
C. School Administrator Characteristics							
12. Male (0/1)	0.969	0.174	0.017 (0.028)	-0.014 (0.032)	-0.000 (0.028)	260	0.784
13. Age (years)	40.307	7.241	1.675 (1.365)	0.592 (1.367)	1.011 (1.387)	260	0.661
14. Higher Education Degree (0/1)	0.846	0.364	-0.022 (0.066)	-0.013 (0.066)	0.019 (0.065)	260	0.920
15. Administration Experience (years)	8.488	6.167	1.216 (1.245)	-0.524 (1.120)	0.935 (1.227)	260	0.487

NOTES. Data source: authors' survey. Table uses full baseline sample. First two columns show the mean and standard deviation of the baseline characteristic in the comparison (small grant, no incentives) cell. Hemoglobin concentration and anemia statistics are shown for the sample of children randomly selected for hemoglobin tests. The remaining characteristics in Panel A are shown for all children in the full sample regardless of whether they were selected for the hemoglobin test. A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). Columns 3 through 5 show coefficients and standard errors (in parentheses) from a regression of the characteristic on incentive group dummy variables estimated using equation (5.3) but controlling only for randomization strata fixed effects. The final column shows the p-value from a Wald test that coefficients are jointly zero. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 5.3: Descriptive Statistics and Balance: Anemia, Test and Dual Incentives (Full Sample)

	Control		Coefficient (standard error) on:			N	P-value: Equality of All Groups
	Mean	SD	Anemia	Test	Dual		
A: Child Characteristics							
1. Hemoglobin Concentration (g/L)	134.191	12.912	-0.974 (0.970)	-1.528 (1.027)	0.515 (0.880)	6456	0.215
2. Anemic (0/1)	0.233	0.423	0.015 (0.018)	0.008 (0.017)	0.002 (0.015)	6456	0.841
3. Baseline Exam Score (Normalized)	0.015	0.975	-0.011 (0.030)	-0.022 (0.030)	-0.067** (0.030)	9131	0.143
4. Age (years)	10.713	1.173	-0.072 (0.105)	-0.007 (0.103)	-0.120 (0.102)	9135	0.581
5. 5th Grade (0/1)	0.531	0.499	-0.018 (0.013)	-0.012 (0.015)	0.006 (0.014)	9135	0.291
6. Female (0/1)	0.485	0.500	0.005 (0.015)	0.002 (0.016)	0.023 (0.016)	9134	0.458
B: School Characteristics							
7. Number of Students	207.094	64.823	0.795 (14.156)	-13.624 (13.603)	4.600 (14.242)	130	0.496
8. Has Kitchen (0/1)	0.063	0.246	0.097 (0.079)	0.083 (0.093)	0.089 (0.074)	130	0.531
9. Student-Teacher Ratio	16.228	4.227	0.600 (1.152)	0.544 (1.132)	2.479** (1.163)	130	0.188
10. Time to Furthest Village Served (mins)	62.031	36.695	-3.399 (10.789)	-4.229 (10.548)	3.939 (10.073)	130	0.896
11. Percent Boarding Students (%)	5.327	11.404	0.879 (2.842)	-1.585 (3.278)	-1.393 (2.712)	130	0.812
C. School Administrator Characteristics							
12. Male (0/1)	0.938	0.246	0.045 (0.037)	0.014 (0.043)	0.045 (0.038)	130	0.397
13. Age (years)	39.313	7.253	1.933 (1.807)	2.066 (1.845)	3.512** (1.703)	130	0.234
14. Higher Education Degree (0/1)	0.906	0.296	-0.046 (0.080)	-0.117 (0.089)	-0.134 (0.088)	130	0.387
15. Administration Experience (years)	8.031	6.141	1.465 (1.808)	0.349 (1.603)	2.642 (1.692)	130	0.391

NOTES. Data source: authors' survey. Table uses full baseline sample of small block grant schools. First two columns show the mean and standard deviation of the baseline characteristic in the comparison (small grant, no incentives) cell. Hemoglobin concentration and anemia statistics are shown for the sample of children randomly selected for hemoglobin tests. The remaining characteristics in Panel A are shown for all children in the full sample regardless of whether they were selected for the hemoglobin test. A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). Columns 3 through 5 show coefficients and standard errors (in parentheses) from a regression of the characteristic on incentive group dummy variables estimated using equation (5.3) but controlling only for randomization strata fixed effects. The final column shows the p-value from a Wald test that coefficients are jointly zero. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 5.4: Descriptive Statistics and Balance: Anemia, Test and Dual Incentives (Small Grant)

	Control		Coefficient (standard error) on:			N	P-value: Equality of All Groups
	Mean	SD	Anemia	Test	Dual		
A: Child Characteristics							
1. Hemoglobin Concentration (g/L)	133.724	12.701	-0.890	0.238	-0.243	6422	0.740
			(0.926)	(0.895)	(0.879)		
2. Anemic (0/1)	0.232	0.422	0.022	0.007	0.014	6422	0.466
			(0.015)	(0.011)	(0.013)		
3. Baseline Exam Score (Normalized)	-0.016	1.024	0.025	0.019	0.021	8987	0.837
			(0.028)	(0.028)	(0.030)		
4. Age (years)	10.669	1.161	-0.054	-0.112	-0.010	8989	0.540
			(0.077)	(0.080)	(0.092)		
5. 5th Grade (0/1)	0.523	0.500	-0.002	-0.010	0.007	8989	0.680
			(0.015)	(0.014)	(0.016)		
6. Female (0/1)	0.485	0.500	0.003	0.002	-0.008	8989	0.902
			(0.015)	(0.016)	(0.016)		
B: School Characteristics							
7. Number of Students	201.697	58.681	13.309	9.067	9.355	130	0.834
			(14.822)	(13.748)	(14.511)		
8. Has Kitchen (0/1)	0.152	0.364	-0.019	-0.050	-0.012	130	0.946
			(0.096)	(0.087)	(0.092)		
9. Student-Teacher Ratio	15.594	5.030	1.711	0.276	0.964	130	0.430
			(1.130)	(1.171)	(1.263)		
10. Time to Furthest Village Served (mins)	66.667	56.259	-2.484	-0.044	7.518	130	0.930
			(14.490)	(14.479)	(17.367)		
11. Percent Boarding Students (%)	6.482	14.491	-0.825	0.540	1.004	130	0.976
			(3.967)	(3.859)	(4.338)		
C. School Administrator Characteristics							
12. Male (0/1)	1.000	0.000	--	--	--	--	--
			--	--	--		
13. Age (years)	41.270	7.208	0.954	-0.419	-1.160	130	0.742
			(2.062)	(2.098)	(2.012)		
14. Higher Education Degree (0/1)	0.788	0.415	0.005	0.049	0.144	130	0.306
			(0.098)	(0.099)	(0.092)		
15. Administration Experience (years)	8.930	6.255	0.882	-0.608	-0.143	130	0.855
			(1.892)	(1.628)	(1.907)		

NOTES. Data source: authors' survey. Table uses full baseline sample of large block grant schools. First two columns show the mean and standard deviation of the baseline characteristic in the comparison (small grant, no incentives) cell. Hemoglobin concentration and anemia statistics are shown for the sample of children randomly selected for hemoglobin tests. The remaining characteristics in Panel A are shown for all children in the full sample regardless of whether they were selected for the hemoglobin test. A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). Columns 3 through 5 show coefficients and standard errors (in parentheses) from a regression of the characteristic on incentive group dummy variables estimated using equation (5.3) but controlling only for randomization strata fixed effects. The final column shows the p-value from a Wald test that coefficients are jointly zero. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 5.5: Descriptive Statistics and Balance: Anemia, Test and Dual Incentives (Large Grant)

5.4.2 Correlation between Hb Concentration and Exam Scores at Baseline

To interpret results, it is useful to see how the two primary outcomes, Hb concentration and standardized exam scores, are functionally correlated. Figure 5.5 shows the estimated relationship between hemoglobin concentration and student scores on standardized exams (separately for math and Chinese) using data on 14,872 4th and 5th grade students for whom Hb measurements were collected. Hemoglobin concentration in g/L is shown on the x-axis and normalized exam scores are shown on the y-axis. These functional relationships are estimated using restricted cubic spline regressions with 4 knots. The "unadjusted" estimates do not adjust for any other covariates. The "adjusted" estimate adjusts linearly for student age, grade, sex, ethnicity, whether the child boards at school, household size, the father's migration status, father's education, a household asset index, and province.

It is clear from these estimates that there is indeed a strong relationship between hemoglobin concentration and exam scores in the sample. Further, the relationship is approximately linear, at least in Math ($P_{nonlinear} = 0.93$ for Math and $P_{nonlinear} = 0.57$ for Chinese using the adjusted model). Interestingly, Math scores are positively correlated with Hb concentration well past the WHO defined threshold for anemia (120 g/L for this age group).

5.4.3 Attrition and Non-response

Tables 5.6 and 5.7 assess attrition of students between baseline and endline. I define attrition as a missing Hb concentration measure or exam score at endline when one existed for the baseline. I present OLS regressions of an indicator for a missing values at endline on treatment arm dummy variables and interactions (as in Equation (5.3)) with and without baseline characteristics. All regressions control for county and randomization strata dummy variables.

Table 5.6 presents this analysis for anemia incentive groups and cross-cutting block grant groups. Columns (1) and (2) use the sample of all anemic children in the baseline sample and columns (3) and (4) use the full sample of children for whom we took hemoglobin measurements at baseline. I do find some differential attrition between the treatment arms. In all four regressions,

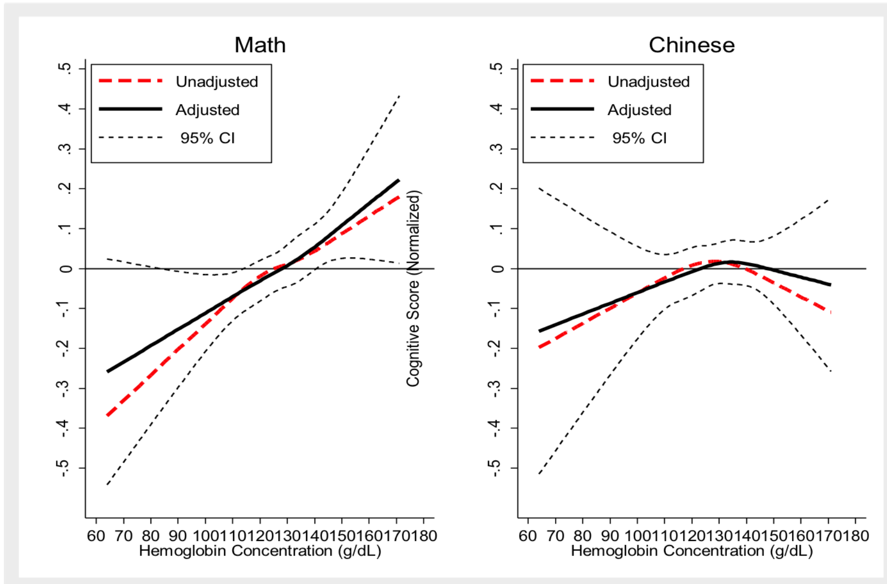


Figure 5.5: Spline Regression Estimates of the Relationship Between Hb and Standardized Scores

the coefficient on the interaction between the small incentive group dummy and large block grant dummy is negative and significant. The level of differential attrition, however, is small in magnitude (the largest estimated difference is 6 percentage points) and the overall amount of attrition in the experiment is small (5.4% in the full sample). Thus, estimates should not be significantly affected. [16]

Table 5.7 shows similar regressions for the large anemia incentive, test and dual incentive groups. Columns (1) and (2) use the full sample, columns (3) and (4) use the sample of schools that received small block grants, and columns (5) and (6) show regressions for large block grant schools. I find no differential attrition across treatment arms for the full sample and subsample of small grant schools. I do, however, find an approximately 3 percentage point increase in attrition due to the test incentive in large grant schools. Compared to the mean in the control (0.078), this is small in magnitude and should not significantly impact results. I will consider the impact of this in interpretation of findings presented in Chapter 6.

[16] In main results (below), I find point estimates for the coefficient on the interaction between the small incentive group and large grant to be close to zero; this small amount of differential attrition is not enough to cause this to be significantly different from zero at a conventional level even with conservative assumptions on the potential outcomes of attriters.

Dependent Variable: Hemoglobin measurement missing at endline				
	Children Anemic at Baseline		Full Sample	
	(1)	(2)	(3)	(4)
A. Treatments and Interactions				
Small Incentive	0.014	0.017	0.016	0.014
	(0.019)	(0.020)	(0.010)	(0.010)
Large Incentive	-0.027	-0.027	0.008	0.007
	(0.017)	(0.017)	(0.009)	(0.009)
Large Grant	0.010	0.009	0.005	0.002
	(0.023)	(0.022)	(0.010)	(0.010)
(Small Incentive)X(Large Grant)	-0.059*	-0.063**	-0.039***	-0.035**
	(0.031)	(0.031)	(0.014)	(0.014)
(Large Incentive)X(Large Grant)	-0.007	0.000	-0.023*	-0.021*
	(0.031)	(0.031)	(0.013)	(0.013)
B. Child Characteristics				
Baseline Hemoglobin Concentration (g/L)	-0.000	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.000)	(0.000)
Age (years)		0.013*		0.016***
		(0.007)		(0.003)
5th Grade (0/1)		-0.011		-0.018***
		(0.011)		(0.006)
Female (0/1)		-0.018		-0.013**
		(0.013)		(0.006)
C. School Characteristics				
Number of Students		-0.000		-0.000
		(0.000)		(0.000)
Has Kitchen (0/1)		0.006		0.009
		(0.025)		(0.012)
Student-Teacher Ratio		-0.001		0.000
		(0.002)		(0.001)
Time to Furthest Village Served (mins)		-0.000		0.000
		(0.000)		(0.000)
Percent Boarding Students (%)		-0.000		-0.000
		(0.001)		(0.000)
"Free Lunch" Policy School		-0.019		-0.036*
		(0.056)		(0.021)
Constant	0.118	0.064	0.064*	-0.047
	(0.103)	(0.138)	(0.036)	(0.058)
Observations	2051	2051	8398	8395
R-squared	0.046	0.051	0.023	0.030
Mean in Small Grant, No Incentive Group	0.073		0.053	

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.1) (in addition to what is shown, controlling for county and randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 5.6: Attrition: Anemia Incentives

Dependent Variable: Test score missing at endline						
	Full Sample		Small Grant Schools		Large Grant Schools	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Treatments						
Large Anemia Incentive	-0.002	-0.001	0.003	0.002	-0.015	-0.009
	(0.007)	(0.007)	(0.010)	(0.009)	(0.011)	(0.011)
Test Incentive	0.011	0.011	0.004	-0.004	0.019	0.027**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.012)	(0.011)
Dual Incentive	-0.005	-0.003	-0.001	-0.000	-0.008	-0.001
	(0.006)	(0.007)	(0.010)	(0.009)	(0.011)	(0.010)
B. Child Characteristics						
Baseline Hemoglobin Concentration (g/L)	0.000	-0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Baseline Standardized Exam Score	-0.009***	-0.006**	-0.008**	-0.004	-0.009**	-0.007
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Age (years)		0.016***		0.017***		0.014***
		(0.003)		(0.004)		(0.004)
5th Grade (0/1)		-0.023***		-0.026***		-0.019**
		(0.005)		(0.007)		(0.008)
Female (0/1)		-0.014***		-0.010		-0.015**
		(0.005)		(0.007)		(0.007)
C. School Characteristics						
Number of Students		-0.000		-0.000**		-0.000
		(0.000)		(0.000)		(0.000)
Has Kitchen (0/1)		0.007		0.012		0.032**
		(0.011)		(0.010)		(0.015)
Student-Teacher Ratio		0.000		-0.000		0.001
		(0.001)		(0.001)		(0.001)
Time to Furthest Village Served (mins)		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)
Percent Boarding Students (%)		-0.001***		-0.000*		-0.002***
		(0.000)		(0.000)		(0.000)
"Free Lunch" Policy School		-0.078***		-0.065***		--
		(0.021)		(0.022)		--
Constant	0.037	-0.013	0.077*	-0.001	-0.000	-0.122*
	(0.036)	(0.053)	(0.042)	(0.068)	(0.058)	(0.069)
Observations	12872	12872	6452	6452	6420	6420
R-squared	0.022	0.029	0.017	0.025	0.041	0.049
Mean in Control (No Incentive) Group	0.063		0.048		0.078	

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables obtained by estimating equation (5.3) (in addition to what is shown, controlling for county and randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 5.7: Attrition: Anemia, Test and Dual Incentives

Chapter 6: Results

6.1 Incentives for Anemia Reduction

6.1.1 Main Effects on Hb Concentration and Anemia Status

Table 6.1 presents results for the effect of principal incentives for anemia reduction on hemoglobin concentration and anemia status. In line with the pre-analysis plan, I focus on results for the subsample of children anemic at baseline (the first two columns of Table 6.1). I also present results for the full sample of children randomly sampled for hemoglobin testing (Columns 3 and 4 of Table 6.1). Rows 1-5 in the table show the coefficients on treatment variables from regressions estimated using Equation (5.3). Row 8 gives the mean in the small block grant, no incentive group (the comparison group).

In addition to Table 6.1, I also present results graphically in Figures 6.1 & 6.2. These figures plot the estimated effects and 95% confidence intervals for each group based on estimated coefficients using the sample of initially anemic students (Columns 1 & 2 in Table 6.1). Figure 6.1 does this in terms of predicted endline hemoglobin concentration and Figure 6.2 in terms of the predicted proportion of anemic students at endline.

	Children Anemic at Baseline		Full Sample	
	Hemoglobin Concentration (g/L)	Anemic at Endline	Hemoglobin Concentration (g/L)	Anemic at Endline
	(1)	(2)	(3)	(4)
1. Small Incentive	-0.387	-0.012	1.055	-0.028
	(1.101)	(0.040)	(0.987)	(0.020)
2. Large Incentive	2.567**	-0.138***	0.918	-0.045**
	(1.044)	(0.039)	(0.946)	(0.022)
3. Large Grant	4.205***	-0.145***	2.871***	-0.073***
	(1.123)	(0.038)	(0.989)	(0.021)
4. (Small Incentive)X(Large Grant)	1.445	-0.042	-0.859	0.027
	(1.541)	(0.056)	(1.340)	(0.027)
5. (Large Incentive)X(Large Grant)	-4.580***	0.196***	-3.304**	0.086***
	(1.586)	(0.058)	(1.404)	(0.031)
6. Observations	1923	1923	7943	7943
7. R-squared	0.303	0.110	0.348	0.120
8. Mean in Small Grant, No Incentive Group	129.900	0.360	136.330	0.180

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.1: Effects of Anemia Reduction Incentives and Block Grant Size on Student Hemoglobin Concentration and Anemia Prevalence

Columns 1 & 2 of Table 6.1 show that the addition of the large incentive to the small block grant had a significant impact on student hemoglobin concentration and anemia rates among students found to be anemic at baseline. The estimated coefficient on the dummy variable for the large incentive group (coefficient β_2 in equation 5.3) implies an average increase in hemoglobin concentration among these children by 2.57 g/L and reduction in anemia prevalence at endline by approximately 14 percentage points (a 39% reduction based on the comparison group anemia rate of 36%). Both effects are significant at 5% and represent meaningful improvements in child health. In contrast, the small incentive had no detectable effect on average hemoglobin concentration or anemia. We can reject that the two coefficients are equal ($\beta_1 = \beta_2$) with a p-value <0.001. Note that while the number of schools in the small incentive group is less than that in the other treatment arms, the standard errors on these coefficients are similar.

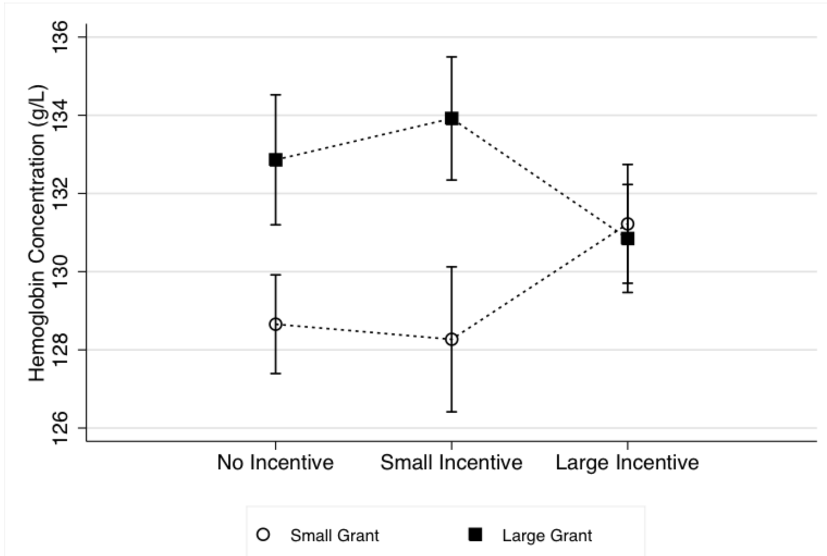


Figure 6.1: Total Effects of Incentives and Block Grants on Student Hemoglobin Concentration (g/L)

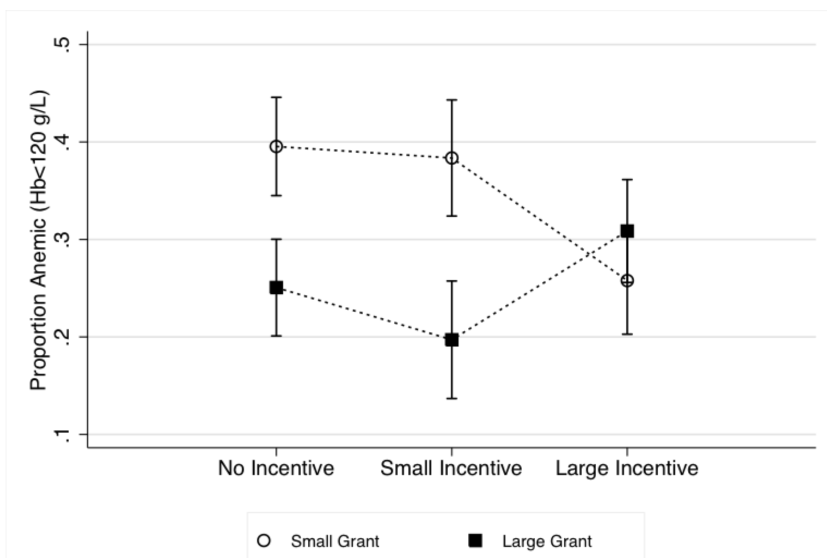


Figure 6.2: Total Effects of Incentives and Block Grants on Student Anemia Prevalence (Hb≤120 g/L)

6.1.2 Mechanisms

The results presented thus far show that large incentives, but not small incentives, produced significant and meaningful reductions in anemia rates among sample students. In this section, I explore avenues through which incentives may have worked by examining student receipts of iron-containing supplements and iron-rich foods, information provided by schools to children and households.

6.1.2.1 Receipts of Supplements and Iron-rich Food

I present results for child and household reported receipt of supplements and iron-rich food in Tables 6.2 & 6.3. Table 6.2 uses the sample of children anemic at baseline and Table 6.3 uses the full hemoglobin sample. Panel A shows secondary outcome variables describing supplement distribution (from household and student surveys) and Panel B shows child reported consumption of food that is (potentially) rich in iron or promotes iron absorption. Information on food consumption was collected using a seven-day recall "food frequency questionnaires" (FFQs) filled by students as part of the endline survey. These questionnaires asked children the number of times they had eaten each of 33 food items in the past seven days, separately for school and home.[1] The final row shows results for the summary index of all variables in the table (using the approach in Anderson (2008)).

[1]Food frequency questionnaires (FFQs) such as the one we use have long been used in nutrition research and have been recommended for use in large surveys of children given low cost and low respondent burden (McPherson et al. 2000; Magarey et al. 2011). While FFQs are generally believed to overestimate consumption relative to other approaches (McPherson et al. 2000), it has been shown that the FFQ responses of children (similar in age to those in our study) regarding their own consumption was more accurate than the responses of their parents (Burrows et al. 2013). Further, previous studies have shown that FFQ responses in China are sensibly correlated with disease outcomes (Zhao et al. 2002).

The results for receipt of supplements and food among children anemic at baseline are consistent with those for hemoglobin (Table 6.2). This is most clearly seen in the results for the summary index (Row 13). The addition of the small incentive to the small grant motivated administrators enough to affect the distribution of supplements (Panel A, Column 2). The large incentive, on the other hand, affected both supplement distribution and children's diets at home. Households with children in large incentive schools report receiving supplements from the school (Table 6.2, Row 1) and children report more frequent consumption of meat, vegetables and fruit at home (Table 6.2, Rows 10-12). This suggests that, when given only a small grant, large incentives led administrators to engage households to improve child nutrition (effectively mobilizing extra-marginal resources). Rural school administrators in China may be particularly effective in this regard given their status in the community and connection to parents. Depending on program goals, however, shifting a portion of this burden to households may not be a desirable result of performance incentives.

These results for children anemic at baseline are largely consistent with those for the full sample (Table 6.3). One exception is that small incentives appear to have had an effect on feeding at home when children not anemic at endline are included. A possible reason for this is that household of non-anemic children may have been more responsive nutritional information provided by principals. Thus, less effort was required to change their feeding behavior.

6.1.2.2 Communication with Households

Tables 6.4 and 6.5 examine more closely how administrators may have worked through households to improve nutrition at home. Here it is also clear that adding large incentives led administrators to engage more with households. Households in large incentive schools report more interactions with school personnel and are more likely to report receiving nutrition related information from the school. Here the results for the sample of children anemic at baseline (Table 6.4) and for the full sample (Table 6.5) are nearly identical suggesting that differences in effects on home feeding (Tables 6.2 and 6.3) are more likely due to differences in household responsiveness

	Mean in Small Grant, No Incentive Group	Coefficient (standard error) on: Small Incentive	Large Incentive	Large Grant	(Small Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	N
A. Iron Supplements							
1. Household received supplements to give to child (Household Response)	0.500	0.038 (0.100)	0.260*** (0.086)	0.101 (0.092)	-0.059 (0.142)	-0.381*** (0.138)	1488
2. School provided supplements to children (Child Response)	0.840	0.200*** (0.065)	0.180*** (0.061)	0.191** (0.075)	-0.444*** (0.106)	-0.332*** (0.093)	1900
3. Times per week supplements distributed by school (Child Response)	2.770	0.697** (0.296)	0.218 (0.304)	0.308 (0.306)	-1.182** (0.480)	-0.465 (0.439)	1911
4. School provided supplements to take home over the weekend (Child Response)	0.150	-0.024 (0.075)	-0.085 (0.057)	0.042 (0.066)	-0.039 (0.081)	0.028 (0.080)	1901
5. All classmates take supplements (Child Response)	0.570	0.202*** (0.068)	0.112 (0.068)	0.185** (0.075)	-0.103 (0.105)	-0.113 (0.099)	1833
6. Days given supplements last month (Child Response)	8.940	0.087 (1.782)	-0.374 (1.760)	2.688 (1.867)	1.006 (2.739)	0.206 (2.518)	1911
B. Food Consumption (Child Food Frequency Questionnaire)							
7. Times consumed meat at SCHOOL in past week	0.580	0.105 (0.216)	0.329 (0.213)	0.227 (0.235)	0.133 (0.317)	-0.706** (0.294)	1923
8. Times consumed vegetables at SCHOOL in past week	1.270	-0.675* (0.346)	0.410 (0.313)	0.834** (0.393)	-0.105 (0.563)	-1.454*** (0.545)	1923
9. Times consumed fruit at SCHOOL in past week	1.300	-0.426 (0.345)	0.275 (0.313)	1.015** (0.463)	-0.858 (0.554)	-1.298** (0.581)	1923
10. Times consumed meat at HOME in past week	3.830	0.427 (0.402)	1.119*** (0.363)	1.035*** (0.394)	-1.045* (0.571)	-1.622*** (0.597)	1923
11. Times consumed vegetables at HOME in past week	11.500	0.556 (0.702)	1.387* (0.708)	1.580* (0.837)	-1.200 (1.123)	-1.736 (1.140)	1923
12. Times consumed fruit at HOME in past week	7.390	0.535 (0.562)	1.037* (0.567)	1.058 (0.657)	-0.942 (0.971)	-2.212** (0.897)	1923
13. Summary Index	-0.070	0.139*** (0.052)	0.166*** (0.053)	0.259*** (0.052)	-0.263*** (0.077)	-0.334*** (0.076)	1923

NOTES. Data source: authors' survey. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration¡120 g/L) at baseline. Dependent variables are shown at left. The final row shows estimates from a regression with an index summarizing all other variables as the dependent variable. This summary index was computed using the method discussed in Anderson (2008). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.2: Child and Household Reported Receipt of Supplements and Iron-Rich Food (Anemic Sample)

	Mean in Small Grant, No Incentive Group	Coefficient (standard error) on: Small Incentive	Large Incentive	Large Grant	(Small Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	N
A. Iron Supplements							
1. Household received supplements to give to child (Household Response)	0.430	0.047	0.213***	0.160**	-0.167	-0.271**	6235
		(0.092)	(0.066)	(0.077)	(0.120)	(0.109)	
2. School provided supplements to children (Child Response)	0.800	0.168***	0.128*	0.152**	-0.402***	-0.263***	7825
		(0.060)	(0.065)	(0.064)	(0.098)	(0.088)	
3. Times per week supplements distributed by school (Child Response)	2.570	0.585**	0.344	0.427	-1.227***	-0.547	7874
		(0.295)	(0.331)	(0.288)	(0.424)	(0.436)	
4. School provided supplements to take home over the weekend (Child Response)	0.130	0.099*	-0.023	0.071	-0.192***	-0.033	7801
		(0.057)	(0.047)	(0.048)	(0.069)	(0.067)	
5. All classmates take supplements (Child Response)	0.690	0.113**	0.084*	0.046	-0.001	-0.002	7622
		(0.053)	(0.047)	(0.052)	(0.079)	(0.071)	
6. Days given supplements last month (Child Response)	6.860	2.164	1.567	4.364**	-1.992	-2.066	7829
		(1.553)	(1.795)	(1.747)	(2.488)	(2.598)	
B. Food Consumption (Child Food Frequency Questionnaire)							
7. Times consumed meat at SCHOOL in past week	0.820	0.191	0.531**	0.124	-0.050	-0.547	7943
		(0.246)	(0.249)	(0.288)	(0.358)	(0.397)	
8. Times consumed vegetables at SCHOOL in past week	0.980	-0.093	0.434	0.442	-0.217	-0.556	7943
		(0.346)	(0.348)	(0.387)	(0.497)	(0.517)	
9. Times consumed fruit at SCHOOL in past week	1.110	-0.093	0.042	0.796***	-0.792**	-0.835**	7943
		(0.229)	(0.217)	(0.304)	(0.398)	(0.374)	
10. Times consumed meat at HOME in past week	4.100	0.712**	0.728***	0.571*	-0.657	-0.615	7943
		(0.304)	(0.270)	(0.308)	(0.449)	(0.456)	
11. Times consumed vegetables at HOME in past week	12.280	0.895*	0.743	0.084	-0.188	0.535	7943
		(0.496)	(0.471)	(0.528)	(0.756)	(0.750)	
12. Times consumed fruit at HOME in past week	7.570	0.992**	0.825**	1.085**	-1.016	-1.373**	7943
		(0.396)	(0.399)	(0.439)	(0.636)	(0.580)	
13. Summary Index	-0.080	0.175***	0.152***	0.193***	-0.283***	-0.208***	7943
		(0.043)	(0.048)	(0.043)	(0.063)	(0.067)	

NOTES. Data source: authors' survey. Table uses full sample. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Dependent variables are shown at left. The final row shows estimates from a regression with an index summarizing all other variables as the dependent variable. This summary index was computed using the method discussed in Anderson (2008). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.3: Child and Household Reported Receipt of Supplements and Iron-Rich Food (Full Sample)

to information rather than targeting by administrators.

6.1.3 Heterogeneous Effects by Administrator and System Characteristics

Tables 6.6 (and continued in 6.7) show how impacts of incentives on hemoglobin concentration vary by characteristics and beliefs of school administrators.[2] For each variable measured at baseline (at left), I estimate Equation (5.3) adding this variable and interaction terms between this variable and β_1 through β_5.

There are three main findings of note. First, both incentives and resources were more effective when administrators had a longer tenure at the school (Row 4 in Table 6.6). Given that the same type of heterogeneity is not found for administration experience generally (Row 3), this could reflect better specific knowledge of the local environment or closer relationships with parents that increase administrators' returns to effort (real and/or perceived). More generally, this suggests that harnessing specific knowledge may be an important channel for the effects of outcome-based incentives.

Second, I find that the effect of large incentives is larger among principals who have a high subjective probability of promotion in the next year (Rows 5-8 in Table 6.6). School systems in China have set schedules for promotion review for school administrators (typically every 3 years). Given this, administrators' subjective expectation of promotion likely reflects whether they do or do not have an upcoming review. A plausible interpretation is therefore that the (experimental) performance incentives may have worked to counter the existing incentives inherent in this review process. That large incentives, but not small incentives, had an effect among this group suggests that large incentives are needed in this context to counter existing 'bureaucratic' incentives.

Third, incentives tend to be less effective if administrators believed households had most responsibility for child health or households were more effective in improving child health (Table 6.7, Rows 10 and 11). These results suggest the importance of administrator perceptions of whether schools have a role in student nutrition. In general, administrators did not take this view: at

[2] For brevity, I only present results for children anemic at baseline. Results for the full sample are qualitatively similar.

	Mean in Small Grant, No Incentive Group	Coefficient (standard error) on:					N
		Small Incentive	Large Incentive	Large Grant	(Small Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	
1. Number of school-wide parent meetings attended this semester	1.440	0.019	0.021	0.676***	-0.978***	-0.682**	1357
		(0.207)	(0.198)	(0.206)	(0.301)	(0.286)	
2. Number of individual meetings with teacher or principal this semester	0.870	0.110	0.503**	0.660***	-0.735**	-0.855**	1345
		(0.185)	(0.231)	(0.251)	(0.325)	(0.376)	
3. School contacted household about student nutrition this semester	0.430	-0.016	0.118*	0.062	-0.062	-0.140	1455
		(0.077)	(0.066)	(0.095)	(0.124)	(0.126)	
4. Household told to give student foods rich in iron	0.270	0.042	0.115**	0.141**	-0.085	-0.273***	1200
		(0.067)	(0.055)	(0.071)	(0.105)	(0.101)	
5. Parent reports knowing of anemia	0.770	0.055	-0.044	0.017	-0.050	0.037	1473
		(0.046)	(0.043)	(0.047)	(0.069)	(0.066)	
6. Parent correctly identifies foods that can prevent anemia (iron rich foods)	1.770	-0.021	0.295	0.176	-0.018	-0.410	1516
		(0.201)	(0.236)	(0.236)	(0.317)	(0.331)	
7. Summary Index	-0.060	0.043	0.139	0.232**	-0.318**	-0.354**	1377
		(0.085)	(0.086)	(0.116)	(0.152)	(0.150)	

NOTES. Data source: authors' survey. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration¡120 g/L) at baseline. Dependent variables are shown at left. The final row shows estimates from a regression with an index summarizing all other variables as the dependent variable. This summary index was computed using the method discussed in Anderson (2008). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.4: Information to Households (Anemic Sample)

	Mean in Small Grant, No Incentive Group	Coefficient (standard error) on: Micro-Incentive	Large Incentive	Large Grant	(Micro-Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	N
1. Number of school-wide parent meetings attended this semester	1.450	0.139	-0.025	0.610***	-0.901***	-0.421	5717
		(0.174)	(0.176)	(0.177)	(0.290)	(0.277)	
2. Number of individual meetings with teacher or principal this semester	0.990	-0.102	0.348*	0.360*	-0.154	-0.527*	5701
		(0.168)	(0.180)	(0.209)	(0.273)	(0.318)	
3. School contacted household about student nutrition this semester	0.420	-0.036	0.096	0.052	-0.021	-0.064	6151
		(0.067)	(0.064)	(0.075)	(0.098)	(0.103)	
4. Household told to give student foods rich in iron	0.260	0.030	0.128**	0.086	-0.110	-0.120	5095
		(0.058)	(0.056)	(0.060)	(0.084)	(0.085)	
5. Parent reports knowing of anemia	0.750	-0.013	0.009	0.023	-0.002	0.037	6175
		(0.040)	(0.035)	(0.038)	(0.059)	(0.052)	
6. Parent correctly identifies foods that can prevent anemia (iron rich foods)	1.820	-0.019	0.292	0.144	-0.220	-0.319	6335
		(0.153)	(0.196)	(0.184)	(0.261)	(0.279)	
7. Summary Index	-0.070	-0.006	0.130	0.182*	-0.210	-0.191	5815
		(0.076)	(0.084)	(0.097)	(0.132)	(0.142)	

NOTES. Data source: authors' survey. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Dependent variables are shown at left. The final row shows estimates from a regression with an index summarizing all other variables as the dependent variable. This summary index was computed using the method discussed in Anderson (2008). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.5: Information to Households (Full Sample)

Baseline Varaible (VAR):	Mean of VAR in Full Sample	Dependent Variable	Coefficient (standard error) on: Small Incentive	Large Incentive	VAR	(VAR)X(Small Incentive)	(VAR)X(Large Incentive)	N
1. Age (years)	40.690	Hb (g/L)	7.368	7.356	0.041	-0.178	-0.120	1923
			(6.317)	(6.084)	(0.100)	(0.143)	(0.154)	
		Anemic (Hb<120)	-0.327	-0.305	-0.001	0.007	0.004	
			(0.224)	(0.258)	(0.004)	(0.005)	(0.006)	
2. Attended College (0/1)	0.270	Hb (g/L)	0.238	2.062	1.828	-2.046	1.015	1923
			(1.411)	(1.634)	(1.630)	(2.901)	(2.703)	
		Anemic (Hb<120)	-0.028	-0.126**	-0.090	0.008	-0.014	
			(0.050)	(0.054)	(0.061)	(0.104)	(0.097)	
3. Experience as school administrator (years)	9.070	Hb (g/L)	-3.664*	3.611*	-0.065	0.386	-0.126	1923
			(2.209)	(2.035)	(0.137)	(0.261)	(0.193)	
		Anemic (Hb<120)	0.149*	-0.219***	0.001	-0.020**	0.009	
			(0.078)	(0.074)	(0.005)	(0.008)	(0.007)	
4. Tenure at School (years)	5.530	Hb (g/L)	-6.710***	-0.931	-0.791***	1.277***	0.719**	1923
			(2.157)	(1.693)	(0.213)	(0.399)	(0.292)	
		Anemic (Hb<120)	0.281***	-0.034	0.030***	-0.059***	-0.023**	
			(0.076)	(0.064)	(0.008)	(0.013)	(0.010)	
5. Subjective probability of promotion in next year	10.820	Hb (g/L)	0.122	1.802	-0.024	-0.046	0.092	1923
			(1.385)	(1.166)	(0.033)	(0.058)	(0.056)	
		Anemic (Hb<120)	-0.004	-0.078*	0.003**	-0.001	-0.006***	
			(0.051)	(0.044)	(0.001)	(0.002)	(0.002)	
6. Subjective probability of prmotion in next year>0	0.300	Hb (g/L)	-1.535	0.470	-0.027	2.247	7.723***	1923
			(1.246)	(1.072)	(1.412)	(2.573)	(2.163)	
		Anemic (Hb<120)	0.038	-0.042	0.079	-0.134	-0.361***	
			(0.048)	(0.041)	(0.063)	(0.093)	(0.091)	
7. Subjective probability of promotionin next 5 years	18.910	Hb (g/L)	0.316	1.923	-0.003	-0.049	0.042	1923
			(1.413)	(1.235)	(0.024)	(0.039)	(0.044)	
		Anemic (Hb<120)	-0.010	-0.071	0.001	0.000	-0.004**	
			(0.051)	(0.047)	(0.001)	(0.001)	(0.002)	
8. Subjective probability of promotion in next 5 years>0	0.450	Hb (g/L)	-1.667	-0.290	-0.511	2.898	5.969***	1923
			(1.219)	(1.275)	(1.388)	(2.279)	(2.123)	
		Anemic (Hb<120)	0.051	0.005	0.095	-0.150*	-0.318***	
			(0.048)	(0.053)	(0.058)	(0.082)	(0.083)	

NOTES. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration¡120 g/L) at baseline. Dependent variables are shown at left. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.6: Heterogeneous Effects of Anemia Reduction Incentives by Administrator Characteristics

Baseline Varaible (VAR):	Mean of VAR in Full Sample	Dependent Variable	Coefficient (standard error) on: Small Incentive	Large Incentive	VAR	(VAR)X(Small Incentive)	(VAR)X(Large Incentive)	N
9. Believes educators have a responsibility to improve health	0.150	Hb (g/L)	-1.352	2.119*	-1.681	4.480	-0.719	1923
			(1.180)	(1.089)	(1.858)	(2.956)	(2.861)	
		Anemic (Hb<120)	-0.014	-0.154***	-0.016	0.021	0.138	
			(0.047)	(0.044)	(0.067)	(0.110)	(0.112)	
10. Believes parents have most responsibility to improve health	0.680	Hb (g/L)	2.889	1.888	1.876	-4.983**	0.260	1923
			(1.861)	(1.699)	(1.440)	(2.376)	(2.016)	
		Anemic (Hb<120)	-0.081	-0.029	-0.024	0.108	-0.137*	
			(0.072)	(0.071)	(0.061)	(0.092)	(0.080)	
11. Believes parents more effective than school in improving student health	0.750	Hb (g/L)	6.927**	7.581***	3.160	-8.754***	-5.420**	1923
			(2.923)	(2.411)	(1.974)	(3.214)	(2.713)	
		Anemic (Hb<120)	-0.244**	-0.258**	-0.105	0.282**	0.130	
			(0.115)	(0.118)	(0.100)	(0.125)	(0.130)	
12. Believes nutrition influences grades	0.730	Hb (g/L)	4.458**	6.404***	0.074	-6.302**	-5.539**	1923
			(2.239)	(2.180)	(1.670)	(2.510)	(2.544)	
		Anemic (Hb<120)	-0.228***	-0.322***	-0.051	0.282***	0.262**	
			(0.084)	(0.099)	(0.076)	(0.093)	(0.113)	
13. Risk aversion	0.73	Hb (g/L)	-0.900	7.040***	0.449	0.726	-5.900**	1923
			(2.321)	(2.622)	(1.787)	(2.571)	(2.902)	
		Anemic (Hb<120)	-0.057	-0.312***	-0.014	0.052	0.223*	
			(0.074)	(0.102)	(0.070)	(0.086)	(0.115)	
14. Risk aversion > Median	0.49	Hb (g/L)	3.413*	4.971***	2.751*	-6.791***	-4.044*	1923
			(1.916)	(1.789)	(1.611)	(2.220)	(2.393)	
		Anemic (Hb<120)	-0.127**	-0.207***	-0.062	0.223***	0.132	
			(0.058)	(0.061)	(0.058)	(0.071)	(0.091)	
15. Prosocial motivation index	-0.010	Hb (g/L)	-0.273	2.371**	-0.611	-4.860**	-0.192	1923
			(1.065)	(1.030)	(0.810)	(2.112)	(1.328)	
		Anemic (Hb<120)	-0.029	-0.136***	0.020	0.230***	-0.003	
			(0.038)	(0.037)	(0.026)	(0.070)	(0.047)	
16. Prosocial motivation index > Median	0.500	Hb (g/L)	2.991*	2.500	-0.827	-6.581***	-0.644	1923
			(1.629)	(1.764)	(1.639)	(2.123)	(2.371)	
		Anemic (Hb<120)	-0.178***	-0.154**	-0.014	0.310***	0.061	
			(0.064)	(0.068)	(0.065)	(0.079)	(0.093)	
17. Intrinsic motivation index	-0.030	Hb (g/L)	-0.401	2.623**	-0.181	-1.719	0.047	1923
			(1.111)	(1.078)	(0.968)	(1.311)	(1.360)	
		Anemic (Hb<120)	-0.013	-0.136***	0.025	0.104***	-0.010	
			(0.038)	(0.039)	(0.039)	(0.040)	(0.051)	
18. Intrinsic motivation index > Median	0.490	Hb (g/L)	1.072	1.925	-0.320	-3.948*	1.231	1923
			(1.638)	(1.818)	(1.496)	(2.245)	(2.393)	
		Anemic (Hb<120)	-0.036	-0.115*	0.056	0.092	-0.043	
			(0.061)	(0.068)	(0.059)	(0.083)	(0.091)	

NOTES. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration¡120 g/L) at baseline. Dependent variables are shown at left. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.7: Heterogeneous Effects of Anemia Reduction Incentives by Administrator Characteristics (Cont'd)

baseline around 15% of administrators believed that schools have a responsibility for student nutrition and around 68% believed this was the responsibility of parents.

The importance of administrator implicit incentives to focus on traditional school functions (and organizational mission) is also apparent in how the effects of incentives vary with whether an administrator thought of nutrition as a having a significant influence on grades (Table 6.7, Row 12). This survey question gave administrators a list of factors that potentially influence student grades and asked administrators to identify which were most important (they could also write-in if something was not listed). We find that if an administrator did not choose student nutrition, the incentives were more effective. A likely explanation is that incentives were not necessary if principals made a strong enough connection between the tasks of improving nutrition and education (which they have stronger existing incentives to do).

Finally, the impact of incentives (small incentives in particular) varied significantly with baseline survey measures of risk aversion and pro-social and intrinsic motivation.[3] Incentives were less effective among more risk-averse individuals (Table 6.7, Row 13 & 14) which could either have effected how individuals perceived an incentive itself or reflect these individuals being more constrained by their unwillingness to engage in something like distributing supplements to students. Likewise, small incentives were much less effective among more pro-socially motivated individuals (those who report liking their job because they care about helping others—as measured using the psychological scales from Grant (2008)), possibly reflecting a 'crowding-out' effect as is widely discussed in the literature. That is, while the small incentive increases the performance of school principals who score low in our pro-sociality scale, it detracts from those who score highly in the pro-sociality scale. This result coincides with previous studies suggesting that small rewards can be detrimental among intrinsically motivated individuals (Gneezy & Rustichini 2000).

I do not find that the effects of incentives (either large or small) varied significantly with variables related to how school systems manage school principals (existing performance incentives,

[3]Risk aversion was measured with a single stated preference survey question asking administrators how much they would pay for a lottery ticket with a 50-50 chance of winning 1000 yuan. We identify administrators as risk averse if their response is below the median for responses in the sample. While not an ideal measure, this response should provide a rough measure of risk aversion. Our interpretation here assumes that measurement error is classical, but it is possible that it is correlated with unobservable factors.

decision making autonomy and monitoring - Table 6.8).[4]

6.2 Resources and Incentives

6.2.1 Main Effects on Hb Concentration and Anemia Status

Estimates in Table 6.1 show that that additional resources under the control of school administrators (the large block grant) had a large impact on anemia prevalence, absent any explicit incentives. The estimates in columns 1 & 2 of Table 6.1 for the effect of the large block grant (β_3 in equation 1) show that additional resources alone increased mean Hb by 4.2 g/L and reduced anemia prevalence by 14.5 percentage points (a 40% reduction).

It is useful to compare the compare the effect of additional resources to the effect of incentives alone in terms of costs.[5] Although additional resources were more effective in reducing anemia rates than large incentives, they are much more costly. The average realized incentive payout to administrators in the large incentive, small block grant group was 3,303 yuan ($524.29), or 15.95 yuan ($2.50) per child based on the average school size. Increasing the size of the block grant by 0.4 yuan per child per day (for 120 total school days) increase the total size of the grant by 9,936 yuan ($1,577) per school on average. Based on the estimates in column 2 in Table 3 (and endline anemia prevalence in the comparison group), adding the large incentive to the small grant averted 6.57 cases of anemia and adding additional resources averted 6.91 cases per school among children initially anemic. Thus, the cost per case of anemia averted among this group was 503 yuan ($79.80) for the large incentive and 1,438 yuan ($228.24) for the large grant. Using estimates for the full sample in columns 3 and 4 of Table 4 (which are more appropriate for cost effectiveness calculations in this context), we estimate that the large incentive averted 9.32 cases of anemia on average per school and the larger grant averted 15.11 cases. For the full sample, therefore, the

[4] School principals are evaluated as part of the cadre evaluation system (ganbu kaohe zhidu). Under this system, administrators are typically evaluated based on a rubric of pre-defined criteria (Whiting 2004). These criteria and weights placed on each are set at the local level and vary widely by locality. Roughly 50% of schools in the sample had existing performance pay schemes that applied to school administrators. Based on anecdotal evidence from structured interviews, rewards under this "Teacher Performance Pay Policy" do not vary significantly and most variation in rewards is due to differences in seniority or work hours, not student outcomes.

[5] Calculations here are admittedly simplistic. I do not account for monitoring costs or other nutritional benefits, for example (the cost of figner prick Hb assesments is neglegible when taken to scale and could be incorporated into existing wellness checks). My goal is merely to provide an indication of relative cost effectiveness.

Baseline Varaible (VAR):	Mean of VAR in Full Sample	Dependent Variable	Coefficient (standard error) on: Small Incentive	Large Incentive	VAR	(VAR)X(Small Incentive)	(VAR)X(Large Incentive)	N
1. School system has existing pay for performance scheme (0/1)	0.5	Hb (g/L)	-0.641	1.693	-0.220	0.476	1.260	1923
			(1.587)	(1.468)	(1.874)	(2.278)	(2.400)	
		Anemic (Hb<120)	-0.006	-0.121**	-0.075	0.013	-0.016	
			(0.054)	(0.052)	(0.074)	(0.084)	(0.091)	
2. Administrator autonomy andex	0.04	Hb (g/L)	-0.232	2.305**	-0.047	4.009	4.054	1923
			(1.225)	(1.103)	(3.276)	(4.460)	(4.120)	
		Anemic (Hb<120)	-0.011	-0.127***	0.025	-0.067	-0.236	
			(0.045)	(0.041)	(0.129)	(0.167)	(0.158)	
3. Administrator autonomy index>=Median (0/1)	0.61	Hb (g/L)	0.100	1.335	0.098	-0.972	2.077	1923
			(1.549)	(1.507)	(1.584)	(2.746)	(2.419)	
		Anemic (Hb<120)	-0.025	-0.088	0.009	0.015	-0.094	
			(0.065)	(0.056)	(0.059)	(0.102)	(0.091)	
4. Times someone from school district made monitoring visit to schools in past year	8.95	Hb (g/L)	-1.098	0.394	0.223**	-0.009	0.221	1923
			(1.818)	(1.943)	(0.111)	(0.167)	(0.194)	
		Anemic (Hb<120)	0.046	-0.053	-0.007*	-0.006	-0.009	
			(0.060)	(0.075)	(0.004)	(0.006)	(0.008)	
5. Monitoring visits to schools>=Median (0/1)	0.38	Hb (g/L)	-0.079	2.351*	3.625***	-1.425	0.566	1923
			(1.512)	(1.342)	(1.385)	(2.297)	(2.120)	
		Anemic (Hb<120)	-0.004	-0.123***	-0.163***	-0.022	-0.042	
			(0.051)	(0.045)	(0.061)	(0.084)	(0.085)	
6. Times school district had school administrator meetings last year	11.51	Hb (g/L)	2.477	1.317	0.197*	-0.286	0.135	1923
			(2.317)	(2.052)	(0.104)	(0.186)	(0.163)	
		Anemic (Hb<120)	-0.158*	-0.062	-0.008**	0.015**	-0.008	
			(0.083)	(0.074)	(0.003)	(0.006)	(0.006)	
7. School administrator meetings>=Median (0/1)	0.34	Hb (g/L)	0.524	1.978*	2.835*	-3.896	1.903	1923
			(1.318)	(1.145)	(1.535)	(2.952)	(2.088)	
		Anemic (Hb<120)	-0.049	-0.101**	-0.101	0.168*	-0.109	
			(0.048)	(0.043)	(0.063)	(0.101)	(0.081)	

NOTES. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (1) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration<120 g/L) at baseline. Dependent variables are shown at left. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.8: Heterogeneous Effects by School System Characteristics

cost per anemia case averted was 354 yuan ($56.19) for the large incentive and 658 yuan ($104.44) for the larger grant. Of course, incentive payouts would be essentially costless if expected total compensation (base pay and incentive payments) remains constant (for example if incentives are included in subsequent pay increases).

It is also clear from Table 6.1 that incentives and increased resources under managerial control are substitutes. The coefficient on the interaction term between the large incentive and large block grant groups (β_5 in Equation 1) is large and significant in all four specifications (Table 2, Row 5). The combined effect of the large incentive and large block grant together still outperforms the comparison group: $\beta_2 + \beta_3 + \beta_5 = 2.19$ (p-value = 0.023) for Hb concentration and -0.087 (p-value = 0.016) for anemia prevalence.[6] On the other hand, the marginal effect of the large incentive given the large block grant is negative and marginally significant: $\beta_2 + \beta_5 = -2.01$ (p-value = 0.08) for Hb concentration and 0.058 (p-value = 0.141) for anemia prevalence. For the full sample, $\beta_2 + \beta_5 = -2.39$ (p-value = 0.01) for Hb concentration and 0.041 (p-value = 0.041) for anemia prevalence. Thus, the substitution between incentives and additional resources was so large that, in the presence of the large block grant, large incentives appear to have had a crowding-out effect.

6.2.2 Intermediate Outcomes

6.2.2.1 Use of Block Grants

An important distinction between managers and front-line workers is the control they exert over use of an organization's resources. How administrators choose to allocate block grants as a result of incentives may therefore be instructive of the role of incentives for individuals with more decision making authority within service organizations. In the case of small block grants (results presented in Section 6.1) there was limited scope for the use of block grants to affect outcomes given small grant size. Large block grants, however–in addition to increasing the overall level of resources available to administrators–also increase their relative authority over resource allocation

[6] For the full sample, $\beta_2 + \beta_3 + \beta_5 = 0.48$ (p-value = 0.579) for Hb concentration and -0.032 (p-value = 0.085) for anemia prevalence.

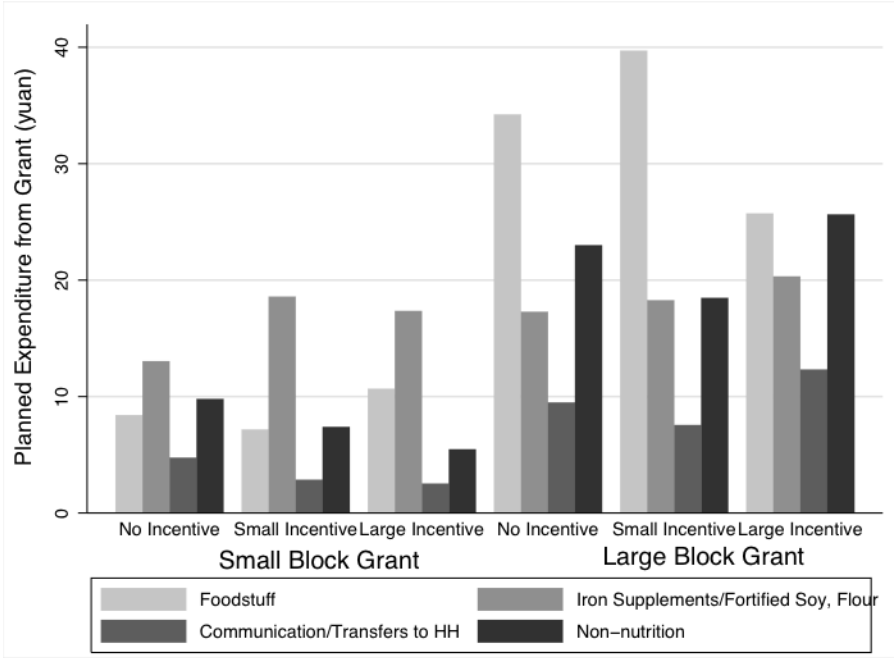


Figure 6.3: Planned Use of Block Grants

and the potential for the allocation of block grant funds to affect outcomes.

Results for the planned and reported use of block grants are shown in Table 6.9 (and graphically in Figures 6.3 and 6.4). Planned block grant use amounts (Panel A) are from a non-binding budget filled by administrators immediately after they learned of incentives and the grants they would receive. Reported block grant use (Panel B) is from an identical budget from filled by administrators at the time of the endline survey.[7] In this table, the reported expenditure per student on each of four categories is regressed on treatment dummies and interactions as in Equation (5.3). Due to limited power for regressions at the school level, however, these regressions do not include county and randomization strata dummies. [8]

[7]Although reported use of block grants by administrators is likely subject to some reporting bias, this would only strengthen the results that I find.

[8]This is a deviation from the pre-analysis plan. Including controls as pre-specified subsumes degrees of freedom and reduces the efficiency of school-level regressions.

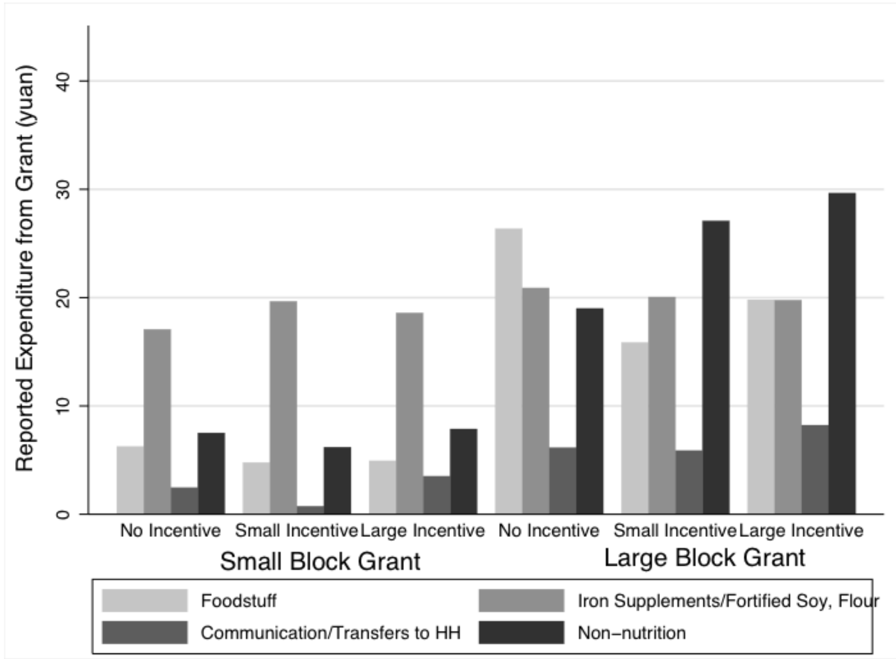


Figure 6.4: Reported Use of Block Grants

The four spending categories that I examine are foodstuff, supplements/fortified flour and soy, communications and transfers to households, and those not explicitly related to nutrition. "Foodstuff" includes all school purchases of food (regardless of iron content). "Supplements/fortified flour and soy" includes school spending on supplements containing iron and fortified flour and soy. Higher spending in this category may suggest an approach of school administrators that was more narrowly focused on iron delivery. "Communications/Transfers to households" includes spending on parent meetings (possibly related to nutrition, but, also possibly not related to nutrition), other communication or nutrition related education materials and transfers such as food subsidies given directly to households (direct transfers were rare). The "Non-nutrition" category includes spending on categories unrelated to nutrition (including spending on educational inputs such as computers, desks, chairs, stationary, teacher overtime for tutoring, and other spending identified by administrators as education-related).

Although power to detect school level differences is limited, several patterns can be seen in these estimates. First, administrators used additional funds with the larger block grant primarily to increase spending on food and on non-nutrition items. In both the planned (Panel A) and

	Mean in Small Grant, No Incentive Group	Coefficient (standard error) on: Small Incentive	Large Incentive	Large Grant	(Small Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	N
A. Planned Block Grant Use (Per Student)							
1. Foodstuff	8.410	-1.196 (3.468)	2.994 (3.504)	26.135*** (5.236)	6.566 (8.763)	-10.951 (7.136)	170
2. Supplements/Fortified Flour, Soy	13.040	5.321 (3.494)	4.139 (3.271)	4.225 (3.164)	-3.895 (5.236)	-1.395 (4.725)	170
3. Communication/Transfers to Households	4.753	-1.338 (1.989)	-2.242 (1.675)	4.577* (2.721)	-0.297 (3.815)	5.417 (4.162)	170
4. Non-nutrition	9.793	-2.783 (2.835)	-4.886* (2.593)	13.066*** (3.999)	-2.379 (6.050)	6.921 (5.548)	170
B. Reported Block Grant Use at Endline (Per Student)							
5. Foodstuff	6.262	-3.128 (3.023)	-1.780 (2.605)	18.786*** (4.144)	-6.663 (7.018)	-3.424 (5.747)	170
6. Supplements/Fortified Flour, Soy	17.074	2.799 (3.979)	2.289 (3.342)	4.864 (4.047)	-3.730 (6.256)	-4.685 (5.604)	170
7. Communication/Transfers to Households	2.462	-1.494* (0.895)	1.053 (1.192)	3.673*** (1.345)	1.318 (2.510)	1.087 (2.878)	170
8. Non-nutrition	7.513	-0.154 (2.834)	-0.267 (3.141)	10.752** (5.172)	8.525 (11.488)	11.781* (7.007)	170

NOTES. Data source: authors' survey. Table shows school administrator "planned" (Panel A) and "reported" (Panel B) use of block grant funds recorded by having administrators fill out a (non-binding) budget at the onset of the nutrition program and the same budget form reporting actual use of funds at endline. Each row shows coefficient estimates (and robust standard errors) from a seperate regressions estimated using equation (5.3) but excluding county and randomization strata dummies. Dependent variables, the reported expenditure of grants by category (per student), are shown at left. "Foodstuff" includes spending on any type of food purchased using block grant funds; "Vitamins/Fortified Flour, Soy" included purchases of multivitamins or fortified flour or soy sauce; "Communication/Transfers to Households" includes funds used for parent meetings, informational materials for parents, or direct transfers to households (in cash or kind); "Non-nutriton" includes spending on education expenditures teacher training, teacher overtime, tutoring fees for students, academic supplies (computers, stationary, desk chairs, etc) and incentives/prizes for teachers and students. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.9: Effects of Resources and Incentives on Planned and Reported Use of Block Grants

reported (Panel B) budgets, increasing the size of the grant did not significantly increase spending on supplements and fortified soy/flour. This is most likely due to administrators allocating funds to supplements by calculating the number of daily supplements required for the number of students in their school (and not purchasing more than this).

It is also clear from Table 6.9 that administrators with large incentives and small block grants did not allocate expenses across categories differently from those with a small block grant and no incentive (though they initially planned to allocate significantly less to "non-nutrition" purposes). Hence, better anemia results for administrators with large incentives and small grants are likely due to the extra effort that they or those that work for them exerted, and not to allocating the grant differently.

Finally, incentives tied to student anemia reductions, in schools with large budgets, increased expenditure on items unrelated to nutrition. For the large incentive, this is evident in both the planned and reported budgets (but is only significant for reported budgets). While surprising, these results are consistent with the main student nutrition results (above) and student and household reported receipts (below).

6.2.2.2 Receipts of Supplements and Iron-rich Food and Communication with Households

Adding resources with the larger grant increased child receipt of fruits and vegetables at school, consistent with the budget results showing larger expenditure on food (Table 6.2 & 6.3, Rows 8 & 9). Importantly, children and households in large block grant schools were also significantly more likely to report receiving supplements from the schools. Combined with the observation that large block grant schools did not spend significantly more from the block grant on supplements, suggests that the large block grant also led to a more efficient use of inputs. Thus, increasing the amount of resources under administrator control appears to have also led to an increase in effort. This is further supported by the increase in communication with households and student-reported consumption of iron-rich foods at home (Tables 6.2 & 6.3, rows 10-12; Tables 6.4 & 6.5). Increasing resources led administrators to work through households, absent any explicit incentives to do so. It is possible that this reflects motivation from additional autonomy heavily emphasized in the public administration literature (Rose-Ackerman 1986; Perry & Wise 1990). This may also reflect a similar 'obligation' effect whereby administrators felt an obligation to implement programs effectively precisely because they received a significant amount of resources.

Adding incentives to the large grant, however, once again, largely negates these effects. The coefficients on the interaction terms (in columns 5 & 6) show that children in schools with administrators who received incentives in addition to the large grant were less likely to receive supplements or more food at school or home. Administrators with incentives in addition to the large grant also appear to have done little to engage households.

6.2.3 Heterogeneity

I find that results for the impact of additional resources vary significantly by whether administrators believe that the school has responsibility to improve health (Table 6.11, Row 9).

Incentives and resources also tend to be less effective if administrators believed households had most responsibility for child health or households were more effective in improving child health. (Table 6.11, Rows 10 and 11). These results suggest the importance of administrator perceptions of whether schools have a role in student nutrition. In general, administrators did not take this view: at baseline around 15% of administrators believed that schools have a responsibility for student nutrition and around 68% believed this was the responsibility of parents. I do not find however that the effect of additional resources varied significantly with any pre specified baseline school system characteristics (Table 6.12).

I discuss heterogeneity in the interaction between resources and incentives in the next section.

6.2.4 Substition and Crowding-out Hypotheses

6.2.4.1 Explanations for Substitution

Above, I find that adding administrator incentives and increased resources under the administrator's control were substitutes. There are a number of plausible explanations for this substitution. As discussed in section 4.2.2, one explanation is that resources and effort are themselves substitutes. In other words, additional units of resources allocated to anemia reduction will decrease marginal returns to effort. As administrators allocate resources, therefore, they may have decreased effort (which is costly in utility terms). Based on the results on mechanisms above, a meaningful manifestation of this may have been how increasing resources affected administrator effort to engage with households and improve children's diets at home. Engaging households can be thought of as a way to bring in extra-marginal resources; as resources available to administrators increase, incentives to expend effort in this dimension are weakened. This effect could lead to a large substiuttion effect, particularly if home based strategies are more effective than school-based strategies.

Another explanation is that incentives were simply not needed to increase effort when administrators had sufficient resources to work with. It is clear from the results above that – even

Baseline Varaible (VAR):	Mean of VAR in Full Sample	Dependent Variable	Coefficient (standard error) on: VAR	(VAR)X(Small Incentive)	(VAR)X(Large Incentive)	(VAR)X(Large Grant)	(VAR)X(Small Incentive)X (Large Grant)	(VAR)X(Large Incentive)X (Large Grant)	N
1. Age (years)	40.690	Hb (g/L)	0.041	-0.178	-0.120	-0.075	0.089	0.087	1923
			(0.100)	(0.143)	(0.154)	(0.147)	(0.200)	(0.233)	
		Anemic (Hb<120)	-0.001	0.007	0.004	0.006	-0.006	-0.003	
			(0.004)	(0.005)	(0.006)	(0.006)	(0.007)	(0.009)	
2. Attended College (0/1)	0.270	Hb (g/L)	1.828	-2.046	1.015	-1.679	2.555	2.343	1923
			(1.630)	(2.901)	(2.703)	(2.454)	(3.571)	(3.758)	
		Anemic (Hb<120)	-0.090	0.008	-0.014	0.053	-0.028	-0.153	
			(0.061)	(0.104)	(0.097)	(0.086)	(0.133)	(0.130)	
3. Experience as school administrator (years)	9.070	Hb (g/L)	-0.065	0.386	-0.126	0.098	-0.673	0.138	1923
			(0.137)	(0.261)	(0.193)	(0.182)	(0.407)	(0.287)	
		Anemic (Hb<120)	0.001	-0.020**	0.009	0.006	0.045***	-0.017*	
			(0.005)	(0.008)	(0.007)	(0.006)	(0.014)	(0.010)	
4. Tenure at School (years)	5.530	Hb (g/L)	-0.791***	1.277***	0.719**	0.683***	-1.256**	-0.814**	1923
			(0.213)	(0.399)	(0.292)	(0.233)	(0.492)	(0.407)	
		Anemic (Hb<120)	0.030***	-0.059***	-0.023**	-0.025***	0.071***	0.025*	
			(0.008)	(0.013)	(0.010)	(0.009)	(0.018)	(0.014)	
5. Subjective probability of promotion in next year	10.820	Hb (g/L)	-0.024	-0.046	0.092	0.036	0.107	-0.068	1923
			(0.033)	(0.058)	(0.056)	(0.048)	(0.086)	(0.073)	
		Anemic (Hb<120)	0.003**	-0.001	-0.006***	-0.005***	0.001	0.007***	
			(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	
6. Subjective probability of prmotion in next year>0	0.300	Hb (g/L)	-0.027	2.247	7.723***	-0.208	1.369	-7.450**	1923
			(1.412)	(2.573)	(2.163)	(2.213)	(3.788)	(3.134)	
		Anemic (Hb<120)	0.079	-0.134	-0.361***	-0.143	-0.021	0.443***	
			(0.063)	(0.093)	(0.091)	(0.090)	(0.135)	(0.120)	
7. Subjective probability of promotionin next 5 years	18.910	Hb (g/L)	-0.003	-0.049	0.042	-0.006	0.093	0.002	1923
			(0.024)	(0.039)	(0.044)	(0.040)	(0.067)	(0.062)	
		Anemic (Hb<120)	0.001	0.000	-0.004**	-0.003**	0.000	0.005**	
			(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	
8. Subjective probability of promotion in next 5 years>0	0.450	Hb (g/L)	-0.511	2.898	5.969***	-0.448	-1.437	-3.664	1923
			(1.388)	(2.279)	(2.123)	(2.193)	(3.269)	(3.150)	
		Anemic (Hb<120)	0.095	-0.150*	-0.318***	-0.163**	0.119	0.325***	
			(0.058)	(0.082)	(0.083)	(0.079)	(0.121)	(0.109)	

NOTES. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration¡120 g/L) at baseline. Dependent variables are shown at left. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.10: Heterogeneous Effects of Resources and Incentives by Administrator Characteristics

Baseline Varaible (VAR):	Mean of VAR in Full Sample	Dependent Variable	Coefficient (standard error) on: VAR	(VAR)X(Small Incentive)	(VAR)X(Large Incentive)	(VAR)X(Large Grant)	(VAR)X(Small Incentive)X (Large Grant)	(VAR)X(Large Incentive)X (Large Grant)	N
9. Believes educators have a responsibility to improve health	0.150	Hb (g/L)	-1.681	4.480	-0.719	7.311**	-5.686	-1.664	1923
			(1.858)	(2.956)	(2.861)	(2.882)	(4.753)	(3.929)	
		Anemic (Hb<120)	-0.016	0.021	0.138	-0.090	0.071	-0.167	
			(0.067)	(0.110)	(0.112)	(0.099)	(0.177)	(0.156)	
10. Believes parents have most responsibility to improve health	0.680	Hb (g/L)	1.876	-4.983**	0.260	-4.094*	6.578*	-3.102	1923
			(1.440)	(2.376)	(2.016)	(2.205)	(3.392)	(2.863)	
		Anemic (Hb<120)	-0.024	0.108	-0.137*	0.069	-0.199	0.229**	
			(0.061)	(0.092)	(0.080)	(0.080)	(0.127)	(0.111)	
11. Believes parents more effective than school in improving student health	0.750	Hb (g/L)	3.160	-8.754***	-5.420**	-4.588*	7.706**	8.356**	1923
			(1.974)	(3.214)	(2.713)	(2.337)	(3.833)	(3.222)	
		Anemic (Hb<120)	-0.105	0.282**	0.130	0.114	-0.259*	-0.163	
			(0.100)	(0.125)	(0.130)	(0.108)	(0.146)	(0.145)	
12. Believes nutrition influences grades	0.730	Hb (g/L)	0.074	-6.302**	-5.539**	-3.633	12.915***	10.190***	1923
			(1.670)	(2.510)	(2.544)	(2.468)	(3.801)	(3.390)	
		Anemic (Hb<120)	-0.051	0.282***	0.262**	0.090	-0.417***	-0.350**	
			(0.076)	(0.093)	(0.113)	(0.096)	(0.143)	(0.142)	
13. Risk aversion	0.73	Hb (g/L)	0.449	0.726	-5.900**	1.636	-4.204	-0.468	1923
			(1.787)	(2.571)	(2.902)	(2.161)	(3.718)	(3.376)	
		Anemic (Hb<120)	-0.014	0.052	0.223*	0.003	0.029	-0.043	
			(0.070)	(0.086)	(0.115)	(0.085)	(0.135)	(0.133)	
14. Risk aversion > Median	0.49	Hb (g/L)	2.751*	-6.791***	-4.044*	-2.537	4.913	2.198	1923
			(1.611)	(2.220)	(2.393)	(2.575)	(3.684)	(3.536)	
		Anemic (Hb<120)	-0.062	0.223***	0.132	0.102	-0.161	-0.119	
			(0.058)	(0.071)	(0.091)	(0.088)	(0.121)	(0.133)	
15. Prosocial motivation index	-0.010	Hb (g/L)	-0.611	-4.860**	-0.192	-0.569	5.819**	2.532	1923
			(0.810)	(2.112)	(1.328)	(1.274)	(2.728)	(1.997)	
		Anemic (Hb<120)	0.020	0.230***	-0.003	0.021	-0.148	-0.043	
			(0.026)	(0.070)	(0.047)	(0.046)	(0.093)	(0.075)	
16. Prosocial motivation index > Median	0.500	Hb (g/L)	-0.827	-6.581***	-0.644	-2.020	9.071***	5.449*	1923
			(1.639)	(2.123)	(2.371)	(2.313)	(3.013)	(3.200)	
		Anemic (Hb<120)	-0.014	0.310***	0.061	0.156*	-0.291**	-0.200	
			(0.065)	(0.079)	(0.093)	(0.085)	(0.117)	(0.123)	
17. Intrinsic motivation index	-0.030	Hb (g/L)	-0.181	-1.719	0.047	0.515	2.112	0.425	1923
			(0.968)	(1.311)	(1.360)	(1.095)	(1.772)	(1.665)	
		Anemic (Hb<120)	0.025	0.104***	-0.010	-0.014	-0.102*	-0.016	
			(0.039)	(0.040)	(0.051)	(0.042)	(0.056)	(0.061)	
18. Intrinsic motivation index > Median	0.490	Hb (g/L)	-0.320	-3.948*	1.231	0.321	3.572	0.511	1923
			(1.496)	(2.245)	(2.393)	(2.075)	(3.632)	(3.153)	
		Anemic (Hb<120)	0.056	0.092	-0.043	-0.092	0.093	0.032	
			(0.059)	(0.083)	(0.091)	(0.076)	(0.128)	(0.119)	

NOTES. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration¡120 g/L) at baseline. Dependent variables are shown at left. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.11: Heterogeneous Effects of Resources and Incentives by Administrator Characteristics (Cont'd)

Baseline Varaible (VAR):	Mean of VAR in Full Sample	Dependent Variable	Coefficient (standard error) on: VAR	(VAR)X(Small Incentive)	(VAR)X(Large Incentive)	(VAR)X(Large Grant)	(VAR)X(Small Incentive)X (Large Grant)	(VAR)X(Large Incentive)X (Large Grant)	N
1. School system has existing pay for performance scheme (0/1)	0.5	Hb (g/L)	-0.220	0.476	1.260	-1.269	3.116	4.944	1923
			(1.874)	(2.278)	(2.400)	(2.497)	(3.480)	(3.830)	
		Anemic (Hb<120)	-0.075	0.013	-0.016	0.017	-0.129	-0.130	
			(0.074)	(0.084)	(0.091)	(0.085)	(0.118)	(0.129)	
2. Administrator autonomy andex	0.04	Hb (g/L)	-0.047	4.009	4.054	-1.144	-4.401	-1.560	1923
			(3.276)	(4.460)	(4.120)	(3.786)	(6.702)	(4.829)	
		Anemic (Hb<120)	0.025	-0.067	-0.236	0.086	-0.036	0.096	
			(0.129)	(0.167)	(0.158)	(0.141)	(0.261)	(0.182)	
3. Administrator autonomy index>=Median (0/1)	0.61	Hb (g/L)	0.098	-0.972	2.077	-1.599	1.959	-1.270	1923
			(1.584)	(2.746)	(2.419)	(2.350)	(3.597)	(3.429)	
		Anemic (Hb<120)	0.009	0.015	-0.094	0.120	-0.184	-0.036	
			(0.059)	(0.102)	(0.091)	(0.076)	(0.131)	(0.121)	
4. Times someone from school district made monitoring visit to schools in past year	8.95	Hb (g/L)	0.223**	-0.009	0.221	0.021	-0.406*	-0.518**	1923
			(0.111)	(0.167)	(0.194)	(0.140)	(0.224)	(0.226)	
		Anemic (Hb<120)	-0.007*	-0.006	-0.009	0.003	0.007	0.015	
			(0.004)	(0.006)	(0.008)	(0.006)	(0.009)	(0.010)	
5. Monitoring visits to schools>=Median (0/1)	0.38	Hb (g/L)	3.625***	-1.425	0.566	0.665	-3.839	-3.177	1923
			(1.385)	(2.297)	(2.120)	(2.204)	(3.031)	(2.911)	
		Anemic (Hb<120)	-0.163***	-0.022	-0.042	0.044	0.063	0.122	
			(0.061)	(0.084)	(0.085)	(0.088)	(0.113)	(0.117)	
6. Times school district had school administrator meetings last year	11.51	Hb (g/L)	0.197*	-0.286	0.135	0.016	0.292	-0.405*	1923
			(0.104)	(0.186)	(0.163)	(0.145)	(0.250)	(0.206)	
		Anemic (Hb<120)	-0.008**	0.015**	-0.008	0.002	-0.017*	0.015**	
			(0.003)	(0.006)	(0.006)	(0.005)	(0.010)	(0.007)	
7. School administrator meetings>=Median (0/1)	0.34	Hb (g/L)	2.835*	-3.896	1.903	-1.314	2.983	-4.724	1923
			(1.535)	(2.952)	(2.088)	(2.040)	(3.983)	(3.042)	
		Anemic (Hb<120)	-0.101	0.168*	-0.109	0.075	-0.235	0.147	
			(0.063)	(0.101)	(0.081)	(0.079)	(0.143)	(0.110)	

NOTES. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). Regressions estimated using only sample of children anemic (altitude adjusted hemoglobin concentration<120 g/L) at baseline. Dependent variables are shown at left. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.12: Heterogeneous Effects of Resources and Incentives by School System Characteristics

without explicit incentives tied to anemia reduction – administrators in large block grant schools did expend more effort than the small block grant group. They both worked to engage with households and appear to have used the vitmains that they purchased more efficiently. It is difficult to say whether this was due to an effective increase in autonomy or to an 'obligation' effect from being endowed with more resources (see section 4.2.1), but if increasing resoruces under administrator control itself strengthened incentives (in the broader sense), adding performance incentives would be less effecive.

6.2.4.2 Explanations for Crowding-out

While the explanations above are likely to explain the bulk of the substitution effect, they cannot account for the full amount of substitution that I find. The degree of substiution was so large that, in the presence of the large grant, incentives had a slight negative effect. This effect was present not only in the results for student anemia outcomes, but also in how principals allocated the block grant and the inputs that students and household report to have received. Although the more standard explanations above can account for resources completely negating the effect of incentives (or vice-versa), they cannot account for this negative effect of incentives. This section presents exploratory analysis of three alternative hypotheses for this 'crowding-out' effect to the extent permitted by the data.

Hypothesis 1: The Combination of Increased Resources and Incentives Convey Information about Task Difficulty Benabou and Tirole (2003) indicate incentives could have a detrimental effect on performance if agents infer that the reason for which the principal set up the incentive is that the task is difficult or unpleasant. In the context of this study, it could be that school administrators perceived the combination of the large grant and large incentive as an indication that addressing student anemia would be difficult (and as a consequence devoted less effort and resource to this task).

Table 6.13 shows estimates for treatment effects on administrator subjective expectations of their ability to affect the anemia rate in their school. Shortly after administrators learned of their contract and block grant, they were told how many children were anemic in their school (from the baseline survey) and asked the following question:

> Think of the number of anemic students found in your school. If you spent (0.3/0.7) yuan per child per day to address anemia, what is the smallest number of children you think would still be anemic at the end of the school year?"

Administrators in the small and large grant groups were told the corresponding amount of money. After answering this question, administrators were asked to provide the largest number they thought would be anemic. Finally, they were asked what percent chance the number of anemic children would be between the mid-point (calculated by the enumerator) and the higher number. Based on responses for the minimum, maximum and percent probability above the mid-point we derived the mean and variance of each administrator's subjective distribution assuming a triangular distribution. [9]As part of the endline survey, we asked a similar series of questions but asked about the number of currently anemic children (administrators had not yet been told the results of endline anemia tests).

The results for responses immediately after treatment are shown in Panel A of Table 6.13 and results for endline questions in Panel B. At both times, administrator's beliefs on the average and variance of the number of anemic children at endline appears unaffected by the large incentive, the large grant or the combination. Interestingly, the results are consistent with the idea that the small incentive conveyed the idea that reducing anemia was relatively easy and hence the expected number of anemic children at endline is smaller. However, this is no longer true if the small incentive is provided together with the large grant (these results also convey the idea that there was informational content in the answers to these questions). While admittedly the power of these tests is low given few school-level observations, the combination of incentives and resources seems unlikely to have conveyed that reducing anemia was a difficult task.

[9]This method of eliciting subjective expectations has been used in a number of studies, including in developing countries (Attanasio & Kaufmann 2009; Delavande et al. 2011).

	Mean in Small Grant, No Incentive Group	Coefficient (standard error) on: Small Incentive	Large Incentive	Large Grant	(Small Incentive)X (LargeGrant)	(Large Incentive)X (Large Grant)	N
A. Immediately after learning of contract							
1. Mean of Ex-post Subjective Distribution	24.710	-8.042**	-3.492	-2.929	9.992**	1.884	170
		(3.549)	(3.409)	(3.003)	(4.244)	(3.956)	
2. Mean of Ex-post Subjective Distribution>=Median	0.410	-0.055	0.174*	0.101	0.055	-0.257*	170
		(0.113)	(0.105)	(0.106)	(0.159)	(0.141)	
3. Variance of Ex-post Subjective Distribution	30.760	-1.161	7.858	3.008	5.979	-23.370	170
		(11.328)	(12.508)	(9.652)	(16.437)	(18.417)	
4. Variance of Ex-post Subjective Distribution>=Median	0.530	0.007	0.011	-0.023	0.113	0.024	170
		(0.094)	(0.073)	(0.072)	(0.126)	(0.105)	
B. Endline Survey							
5. Mean of Ex-post Subjective Distribution	18.990	-3.676	5.084	2.200	-2.346	-6.680	170
		(3.934)	(3.931)	(3.572)	(5.762)	(5.339)	
6. Mean of Ex-post Subjective Distribution>=Median	0.530	-0.167	-0.005	-0.114	0.086	0.067	170
		(0.139)	(0.099)	(0.108)	(0.190)	(0.144)	
7. Variance of Ex-post Subjective Distribution	20.570	-2.410	-10.208	-13.437	-9.081	17.757	170
		(8.002)	(8.785)	(11.239)	(12.716)	(13.376)	
8. Variance of Ex-post Subjective Distribution>=Median	0.500	0.095	0.020	0.079	-0.005	-0.055	170
		(0.140)	(0.105)	(0.124)	(0.204)	(0.166)	

NOTES. Data source: authors' survey. Table shows estimations for the effect of incentive contracts and large grants on the distribution of school administrator subjective expectations over anemia reducations. Panel A shows administrator subjective expectation at baseline (just after learning of their incentives and budget amounts) of how many students would be anemic at the end of the school year. Panel B shows results for a similar question at endline that asked administrator expectations for the current anemia rate (before they learned the result). In each case, administrators were asked the minimum, maximum and percent probability above the median. Mean and variance were derived assuming a triangular distribution. Each row shows coefficient estimates (and robust standard errors) from a seperate regression estimated using equation (5.3). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.13: Administrator Subjective Expectations of Anemia Reduction

Hypothesis 2: More Resources Enable Crowding-out of Pure Intrinsic Motivation   A common criticism of incentives is that they might damage performance if they crowd-out intrinsic motivation (Fehr & Falk 2002; Frey & Jegen 2001; Gneezy et al. 2011; Kamenica 2012; Frey et al. 2013). According to this literature, individuals who are pro-socially or intrinsically motivated (for example, motivated by organizational mission) to exert effort on a task may be less likely to do so when offered external rewards. One prominent theory for crowding out, based on cognitive evaluation theory (CET) in psychology, posits that incentives can be perceived as controlling and reduce an individuals subjective autonomy, crowding out intrinsic motivation as a result (Deci et al. 1999; Frey & Jegen 2001). This type of 'motivational crowding' could thus drive substitution between incentives and resources if, for example, administrators perceived the increase in resources under their control as an increase in their own autonomy. If this was the case, incentives may have been perceived as controlling by this group thus crowding out or displacing intrinsic motivation.

Rows 15-18 in Table 6.11 show heterogeneity in the main results for hemoglobin and anemia by psychological scales for pro-social and intrinsic motivation. Administrators who score low on the scale for pro-social motivation appear to drive a portion of the substitution result between the large incentive and large grant. The substitution result is significantly attenuated for administrators scoring above the median on the prosociality scale in particular (Row 16, next to last column). The amount of substitution does not vary significantly with intrinsic motivation (Rows 17 and 18). To the degree that these scales measure the components of pro-social motivation that would be important in this context, this result suggests that substitution may not be driven by pure pro-social or intrinsic motives. While an important source of internal motivation in an organizational context is a sense of organizational 'mission' (Wilson 1989; Rose-Ackerman 1986), this is plausibly positively correlated with motivation to do one's job for the benefit of others.

Hypothesis 3: Incentive Contracts Change the Nature of the Task   The results for internal motivation in the previous section could suggest that substitution is driven by a more pure pursuit of self-interest. As discussed, one of the strongest incentives faced by school administrators is driven by career concerns. In the Chinese education system, like managers in bureaucracies more generally, career prospects for school administrators are tied asymmetrically to process and outcomes. That is, the career benefits of an attributable good outcome may be more limited in a system where promotion weighted toward tenure given organizational rules are followed. An attributable bad outcome due to failure to adhere to these rules, however, could more easily affect prospects for promotion or even lead to dismissal. A bad outcome could highlight deviation from organizational rules or norms. In China, this asymmetry is heightened by political concerns: a sufficiently bad outcome is much more politically damaging than a good outcome produces political benefits since these benefits tend to be less publicly visible. More generally, a managers career prospects could be affected if he is viewed as acting in a way skewed from the organization's core mission.

In my context, this could help drive substitution between incentives and resources if having a performance incentive makes an individual more responsible—or liable—for a bad outcome. Career concerned and risk averse school administrators may have been cautious in implementing the nutrition program, particularly given heightened concerns for food safety in China (especially for meat).[10] Incentives could heighten this caution if administrators thought incentives shifted liability to them. This was not an issue in the case of the small grant-large incentive arm because administrators were able to do little more than purchase supplements (in their view these were distributed to them by the ministry of education and considered safe). Increasing resources, however, also allowed them to buy food, but they needed to procure this locally on their own. I hypothesize that—if administrators believed that incentives (which were known to their superiors) shifted liability to them—incentives would make them more wary of implementing the nutrition program given uncertainty over food safety. Absent incentives, administrators are just fulfilling their 'obligation' to implement the program but, with incentives, they would be doing it for personal profit. This is consistent with the general insight of Benabou and Tirole (2006) that incentives can affect how one's actions are interpreted. In other words, incentives, particularly if visible to others, can affect an individual's motivation due to the image that an action portrays. Increased resources

[10]In rural primary schools administrators are legally responsible for the wellbeing of children while they are at school. Possibly as a result, administrators with whom we conducted qualitative interviews following the trial commonly identified student safety as a primary burden. For example, one administrator stated, "The safety of students is my responsibility. As long as students are at school, I feel the pressure. Only after I see the school gate locked and students go to sleep can I relax." We also learned of administrators sometimes making decisions through open discussions with teachers (and even holding votes), which could be a strategy for avoiding liability.

with large block grants in this context may have also made administrators' actions more 'visible' (or more likely to be noticed by superiors and others) thus amplifying this crowding-out effect of incentives.

A differnt, yet similar, effect could also occur even if administrators are not overly cautious and administrator actions are not visible. As was shown above, increasing the amount of resources available to administrators to implement anemia reduction programs led to an increase in effort devoted to implementing the program effectively. This increase in effort may have been due to existing institutional incentives (professionalism, obligation, etc) that were reinforced by increasing resources under the control of school administrators. Incentives, however, could have shifted administrators' perception of the task from being part of their professional role (and hence subject to institutional incentives) to one that was outside their professional role. In other words, administrators with incentives felt less of an obligation to effectively implement the program and could pay not to expend effort with forgone rewards. This explanation is similar to Heyman and Ariely (2004) who note that incentives can cause individuals to shift from a social to monetary decision making frame. Here, incentives cause a shift froma professional to personal decision making frame.

I find several results that support mechanisms along these lines. First, substitution is significantly larger among administrators with a non-zero chance of being promoted in the next year (i.e. they will soon be facing a promotion review—Table 6.10, Row 6). While subjective promotion probability could be confounded with unobserved factors (such as motivation and ability), promotion in most school districts is based on a periodic review (typically every 3 years) and, anecdotally, is generally based more on seniority and politics than ability.

Also consistent with this explanation is that the substitution effect is stronger among school administrators who received more visits from school district officials last year, as well as school administrators who had more meetings with their supervisors (Table 6.12, Rows 4 & 6). School administrators who are more closely monitored might be more concerned that superiors would discover a negative outcome (such as a food safety issue). In other words, their actions were more 'visible.'

Thus—while speculative—the available evidence suggests that how incentives change the perception of the task of implementing the programs contributed to the degree of substitution (or crowding-out between incentives and resources) that I find. Administrators could have percieved that incentives changed how others (superiors in particular) viewed the task, particularly perceived liability for bad outcomes. Incentives may have also changed their own perceptions of the task by reframing the task from one that was part of their professional responsibility to one that was not. This is not to say this was the sole factor – it is most plausible that more classical explanations were the major cause. These behavioral and institutional factors nevertheless likely added to the substitution between incentives and resources to the degree that the effect of incentives was negative in the presence of the larger grant.

6.3 Multiple Tasks and Multiple Incentives

6.3.1 Main Effects of Anemia Reduction, Test and Dual Incentives

6.3.1.1 Hemoglobin Concentration and Anemia Prevalence

Tables 6.14 and 6.15 present results for the effects of administrator incentives tied to 1) anemia reduction ("anemia incentives"), 2) standardized exam scores ("test incentives") and 3) both ("dual incentives") on student hemoglobin concentration and anemia prevalence. Table 6.14 shows results for the full ("pooled") sample (including both small and large grant schools). Table 6.15 shows results separately for small and large grant schools. Each table shows the treatment effects for anemia incentives (Row 1), test incentives (Row 2) and how the two incentives interact (Row 3) estimated using equation 5.4 using both the subsample of children anemic at baseline (Hb<120 g/L) and the full sample of students for whom hemoglobin measurements were taken. The tables also give p-values for a test that the coefficients on anemia and test incentives are equal (Row 6) and the estimated effect of the combined (dual) incentive and associated p-value (Rows 7 and 8).

In the pooled sample, both anemia and test incentives had modest effects on anemia preva-

	Children Anemic at Baseline		Full Sample	
	Hemoglobin Concentration (g/L)	Anemic at Endline	Hemoglobin Concentration (g/L)	Anemic at Endline
	(1)	(2)	(3)	(4)
1. Anemia Incentive	0.848	-0.056*	-0.683	0.000
	(0.867)	(0.030)	(0.746)	(0.016)
2. Test Incentive	0.444	-0.058**	-0.210	-0.001
	(0.924)	(0.029)	(0.778)	(0.017)
3. (Anemia)X(Test)	0.243	0.052	1.235	-0.007
	(1.186)	(0.041)	(1.004)	(0.023)
4. Observations	3030	3030	12110	12110
5. R-squared	0.294	0.111	0.332	0.105
6. p-Value [H_0:Anemia = Test]	0.670	0.950	0.480	0.950
7. Dual Incentive	1.540	-0.060	0.340	-0.010
8. p-Value [H_0:Dual = 0]	0.060	0.030	0.650	0.620
9. Control Group (No Incentive) Mean	130.480	0.320	137.200	0.150

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.14: Effects of Anemia and Test Incentives on Hemoglobin and Anemia Status (Pooled)

lence among children found to be anemic at baseline (Table 6.14, Column 2). Both anemia and test incentives reduced the probability that these children were anemic at endline by 5-6 percentage points (Rows 1, 2). These effects are statistically indistinguishable from one another (Row 6). The estimated combined effect of the incentives was similar at 6 percentage points (Row 7). None of the incentives had statistically significant effects on mean Hb concentration among children anemic at baseline nor did they significantly impact Hb concentration or anemia in the full sample of children randomly selected for Hb testing (Columns 3 and 4). There is no evidence of a significant interaction effect between the incentives in the direction of Hb concentration or anemia (Row 3).

Estimating results separately for small and large block grant schools (Table 6.15) reveals substantial heterogeneity in incentive effects between these two groups. In small block grant schools, anemia and test incentives led to mean increases in hemoglobin concentration of approximately 4 g/L among children anemic at baseline and 2-3 g/L among all children with corresponding reductions in anemia prevalence of 19 and 7-8 percentage points respectively (Panel A). Compared to endline anemia rates in the no incentive comparison group, anemia incentives reduced anemia by 54% among initially anemic children and by 41% among all children. Test incentives reduced anemia by 52% among anemic children and by 46% among all children. The effects of anemia and test incentives were statistically similar (Panel A, Row 6). Combining the two incentives had effects comparable in magnitude to each type of incentive alone (Row 7). Although the point estimate for the effect of the dual incentive in the estimation using the subsample of children anemic at baseline is slightly larger in the case of hemoglobin concentration, it is statistically indistinguishable from the coefficients on anemia and test incentives. Coefficients on the interaction between anemia and test incentives show that this substitution in the direction of anemia reduction was statistically significant (Row 3) .

In contrast to school receiving a small block grant, incentives had no effect in large block grant schools. This finding is consistent with results presented in Section 6.2.

Taken together these results for effects on hemoglobin concentration and anemia prevalence suggest that (at least in the case of small block grants) incentives in the direction of test scores are on average well aligned with the task of reducing anemia. That the effects of test incentives on student nutrition were significant and large in magnitude implies that administrators believed that there was a causal relationship between anemia reduction and academic performance (as they were told in the training) and thus viewed anemia reduction as a strategy to increase test scores. Further, strengthening administrator incentives in the direction of test scores (and educating them on the relationship between anemia and academic performance) led to effects comparable in magnitude to incentives for anemia reduction. This finding is somewhat surprising given that test incentives are not perfectly aligned with anemia reduction. Note, however, that test incentives may have been more familiar to school administrators and any weakening of incentive effects may be counterbalanced by more familiarity with the incentive scheme.

	Children Anemic at Baseline		Hb Sample	
	Hemoglobin Concentration (g/L)	Anemic at Endline	Hemoglobin Concentration (g/L)	Anemic at Endline
Panel A: Small Block Grant				
	(1)	(2)	(3)	(4)
1. Anemia Incentive	3.749***	-0.196***	2.359**	-0.073***
	(1.224)	(0.044)	(0.957)	(0.023)
2. Test Incentive	4.161***	-0.188***	3.387***	-0.083***
	(1.077)	(0.042)	(0.836)	(0.020)
3. (Anemia)X(Test)	-3.240*	0.207***	-2.764**	0.101***
	(1.796)	(0.074)	(1.344)	(0.033)
4. Observations	1570	1570	6077	6077
5. R-squared	0.327	0.141	0.380	0.128
6. p-Value [H_0:Anemia = Test]	0.710	0.850	0.270	0.670
7. Dual Incentive	4.670***	-0.180***	2.980***	-0.060***
8. p-Value [H_0:Dual = 0]	<0.01	<0.01	<0.01	<0.01
9. Control Group (No Incentive) Mean	129.900	0.360	136.330	0.180
Panel B: Large Block Grant				
	(1)	(2)	(3)	(4)
1. Anemia Incentive	-0.698	0.012	-1.536	0.013
	(1.202)	(0.042)	(1.014)	(0.022)
2. Test Incentive	-1.667	0.025	-1.592	0.029
	(1.303)	(0.041)	(1.124)	(0.021)
3. (Anemia)X(Test)	3.062**	-0.057	2.446*	-0.042
	(1.481)	(0.053)	(1.335)	(0.028)
4. Observations	1460	1460	6033	6033
5. R-squared	0.329	0.140	0.330	0.116
6. p-Value [H_0:Anemia = Test]	0.530	0.780	0.960	0.510
7. Dual Incentive	0.700	-0.020	-0.680	0.000
8. p-Value [H_0:Dual = 0]	0.470	0.610	0.510	0.980
9. Control Group (No Incentive) Mean	131.070	0.270	138.080	0.130

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.15: Effects of Anemia and Test Incentives on Hemoglobin and Anemia Status (By Grant Size)

Also of note is that incentives in the two directions serve as substitutes to the degree that the combined effect of the incentives (dual incentives) led to no identifiable gain in anemia reduction above each incentive alone. Thus, there is no added benefit of using both incentives simultaneously in terms of anemia reduction. This is despite the fact that the total strength of the dual incentive in this direction is larger than anemia incentives alone.

6.3.1.2 Exam Scores

Using the full sample of students given standardized exams (all 4th and 5th grade students), I find that students in incentive schools (all types) scored slightly lower on endline standardized exams compared to no incentive schools (Table 6.16, Rows, 7-9). While there are no distinguishable effects on Chinese language scores, both anemia and test incentives reduced math scores by approximately 0.05 SD. The magnitude of this effect is slight, however, given standards in the literature.[11]

Within both small and large block grant schools, both types of incentives had little impact on standardized exam scores in the full sample (Table 6.17, Columns 7-9). While estimated (negative) effects among large block grant schools (Panel B) are significant, they are small in magnitude (and comparable to results pooling large and small block grant schools). In small block grant schools (Panel A), there is little evidence that incentives had a significant impact on exam scores on average for the full sample. Point estimates are for the most part, however, are negative.

Among the subsample of children anemic at baseline, I do find some evidence that test incentives actually had a negative effect when the school had a small block grant (Panel A, Row 2, Columns 1-3). These negative effects are meaningful and imply that test incentives reduced standardized exam scores by 0.15 SD in math and 0.13 SD in Chinese.

In general, the lack of a positive effect of incentives on exam scores, particularly among students anemic at baseline, is notable given evidence from previous studies showing iron interventions leading to increases in hemoglobin concentration and reductions in anemia comparable

[11] Note that differential attrition in the test incentive group should not affect results qualitatively. Students scoring higher on baseline exams are less likely to attritt and I find no evidence that treatment effects vary along the baseline distribution of test scores.

	Children Anemic at Baseline			Hemoglobin Sample			Full Sample		
	Pooled	Math	Chinese	Pooled	Math	Chinese	Pooled	Math	Chinese
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Anemia Incentive	-0.045	-0.039	-0.060	-0.025	-0.041	-0.013	-0.033*	-0.053**	-0.013
	(0.037)	(0.052)	(0.047)	(0.021)	(0.029)	(0.026)	(0.018)	(0.025)	(0.021)
2. Test Incentive	-0.069*	-0.075	-0.068	-0.017	-0.043	0.012	-0.028	-0.051*	-0.003
	(0.035)	(0.049)	(0.046)	(0.021)	(0.030)	(0.025)	(0.018)	(0.027)	(0.022)
3. (Anemia)X(Test)	0.068	0.112*	0.043	0.010	0.043	-0.023	0.026	0.066*	-0.016
	(0.049)	(0.067)	(0.061)	(0.029)	(0.041)	(0.037)	(0.026)	(0.037)	(0.032)
4. Observations	3039	1475	1564	12124	5965	6159	17092	8479	8613
5. R-squared	0.599	0.629	0.590	0.579	0.600	0.566	0.581	0.602	0.567
6. p-Value [H_0:Anemia = Test]	0.520	0.490	0.860	0.690	0.930	0.360	0.790	0.960	0.670
7. Dual Incentive	-0.050	0.000	-0.080*	-0.030	-0.040	-0.020	-0.040*	-0.040	-0.030
8. p-Value [H_0:Dual = 0]	0.190	0.960	0.050	0.110	0.140	0.340	0.050	0.140	0.120
9. Control Group (No Incentive) Mean	0.040	0.000	0.090	0.020	0.020	0.020	0.000	0.000	0.000

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline exam scores, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.16: Effects of Anemia and Test Incentives on Math and Chinese Exam Scores (Pooled)

	Children Anemic at Baseline			Hemoglobin Sample			Full Sample		
	Pooled	Math	Chinese	Pooled	Math	Chinese	Pooled	Math	Chinese
Panel A: Small Block Grant									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Anemia Incentive	-0.045	-0.084	-0.009	-0.047	-0.105**	0.009	-0.033	-0.061	-0.013
	(0.069)	(0.097)	(0.083)	(0.030)	(0.044)	(0.037)	(0.026)	(0.038)	(0.032)
2. Test Incentive	-0.126**	-0.145*	-0.134*	-0.034	-0.029	-0.035	-0.024	-0.013	-0.036
	(0.055)	(0.073)	(0.080)	(0.027)	(0.042)	(0.036)	(0.025)	(0.041)	(0.032)
3. (Anemia)X(Test)	0.064	0.101	0.045	0.048	0.085	0.009	0.050	0.065	0.042
	(0.081)	(0.107)	(0.105)	(0.046)	(0.060)	(0.058)	(0.040)	(0.053)	(0.050)
4. Observations	1578	775	803	6105	3000	3105	8658	4296	4362
5. R-squared	0.584	0.629	0.572	0.567	0.596	0.553	0.573	0.602	0.557
6. p-Value [H_0:Anemia = Test]	0.140	0.460	0.140	0.670	0.080*	0.240	0.700	0.180	0.510
7. Dual Incentive	-0.110*	-0.130*	-0.100	-0.030	-0.050	-0.020	-0.010	-0.010	-0.010
8. p-Value [H_0:Dual = 0]	0.05	0.08	0.140	0.290	0.300	0.630	0.820	0.840	0.820
9. Control Group (No Incentive) Mean	0.040	0.020	0.050	-0.010	0.000	-0.010	0.000	-0.010	0.000
Panel B: Large Block Grant									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. Anemia Incentive	-0.055	-0.044	-0.075	-0.035	-0.023	-0.047	-0.054**	-0.070**	-0.033
	(0.045)	(0.059)	(0.058)	(0.029)	(0.039)	(0.035)	(0.024)	(0.033)	(0.030)
2. Test Incentive	-0.026	-0.092	0.001	0.009	-0.042	0.059*	-0.021	-0.066*	0.027
	(0.054)	(0.072)	(0.075)	(0.031)	(0.041)	(0.034)	(0.028)	(0.037)	(0.030)
3. (Anemia)X(Test)	0.128**	0.223**	0.098	0.017	0.043	-0.009	0.041	0.086*	-0.009
	(0.060)	(0.088)	(0.077)	(0.041)	(0.057)	(0.048)	(0.038)	(0.050)	(0.044)
4. Observations	1461	700	761	6019	2965	3054	8434	4183	4251
5. R-squared	0.630	0.665	0.640	0.597	0.617	0.591	0.595	0.612	0.586
6. p-Value [H_0:Anemia = Test]	0.560	0.450	0.260	0.170	0.690	<0.001***	0.240	0.930	0.030**
7. Dual Incentive	0.050	0.090	0.020	-0.010	-0.020	0.000	-0.030	-0.050	-0.010
8. p-Value [H_0:Dual = 0]	0.400	0.180	0.750	0.750	0.630	0.940	0.200	0.170	0.620
9. Control Group (No Incentive) Mean	0.050	-0.030	0.120	0.040	0.040	0.050	0.000	0.010	0.000

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline exam scores, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.17: Effects of Anemia and Test Incentives on Math and Chinese Exam Scores (By Grant Size)

to what I find here also led to significant increases in exam performance on the order of 0.2 SD (e.g. Luo et al. 2012). This lack of a positive effect on exam scores is, however, consistent with results in the pilot study (reported in Sylvia et al. 2013). In the pilot study, I found that incentives for anemia reduction had a slight negative effect on test scores on average but this negative effect was concentrated among students healthy (not anemic) at baseline. Test scores of children that were anemic increased in anemia incentives schools relative to control schools. This pattern of effects could be the result of school inputs being distorted away from "education" inputs (effort and resources that serve as pure academic inputs). Such a distortion could occur through a deliberate redistribution of resources by administrators or through more indirect channels (for example, more teacher attention to previously anemic students in response to an increase in their class participation). In the pilot study, this distortion was not enough to counteract gains from improved health. Here, however, one possibility is that this distortion completely negated gains from improved health.

6.3.2 Heterogeneity by Baseline Quintiles of Hb and Exam Scores

The effect of both types of incentives on the academic performance of students with different initial levels of hemoglobin is likely to vary significantly. Indeed, previous trials that provided iron supplements to children suggest that – for a given amount of iron provided – anemic children experience larger gains in hemoglobin concentrations compared to non-anemic children (Luo et al., 2012; Soemantri, Pollitt & Kim, 1985; Soemantri, 1989). Moreover, the benefits of improved iron status on cognitive function and behavior may not be constant along the distribution of hemoglobin concentration (Sungthong, Mo-suwan & Chongsuvivatwong, 2002). Children who are initially anemic, for example, may benefit more from a given improvement than initially healthy children. Further, administrators may respond to incentives by attempting to target inputs to children who they believe would benefit most from those inputs. [12]

To explore how the impact of the interventions on exam scores varied by baseline anemia

[12] While a possibility in the case of both anemia and test scores, it would likely be difficult to target inputs to anemic children since anemia is not physically apparent except in extreme cases.

status, I estimate how treatment effects vary along the baseline distribution of student hemoglobin concentrations. Specifically, Table 6.18 shows the estimated impacts of incentives separately by quintiles of the baseline hemoglobin distribution. Within each quintile, I estimate impacts on endline hemoglobin concentration (odd-numbered columns) and on standardized math and Chinese exam scores. Panel A shows estimates for small grant schools and Panel B shows estimates using large block grant schools.

Focusing on the sample of small block grant schools within which incentives had significant impacts on hemoglobin concentrations, I find that the effects of anemia incentives and test incentives had more or less consistent effects across the distribution. In the first and second and 5th quintiles of the baseline Hb distribution, anemia incentives increased hemoglobin concentrations by approximately 2.8 g/L. Point estimates for the effect of test incentives are slightly larger, ranging from 2.3 g/L to 4 g/L. This is contrary to what one would expect given that marginal returns to a given amount of iron are decreasing in baseline Hb. Lack of heterogeneity in anemia outcomes may, then, reflect differences in inputs.

Neither incentive had positive effects on exam performance in any quintile. The negative effect of test incentives on academic performance is (somewhat surprisingly) largest among the students at the bottom of the distribution who also saw large gains in hemoglobin concentration. While it is difficult to say what led to this result, one interpretation is that education inputs were reallocated through the two mechanisms alluded to above. First incentives may have led to an overall reallocation from academic inputs to anemia-reduction inputs.[13] Although all students were affected by the reduction in academic inputs, this negative effect on exam scores could have been counteracted by the positive effect of improved health for some students. A second type of reallocation may have occurred due to changes in behavior among initially anemic students whose health was improved due to treatment. For example, if teachers structure classes toward students who are more engaged and "ready to learn" (i.e., allocate their effort where its marginal returns are highest), the improved health of initially anemic students may lead them to demand a larger

[13] Although more likely in anemia incentive schools, this may have also occurred in test incentive schools if administrators overestimated the effect of anemia on academic performance.

	Q1		Q2		Q3		Q4		Q5	
	Hb	Test Score	Hb	Test Score	Hb	Test Score	Hb	Test Score	Hb	Test Score
Panel A: Small Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	2.702*	-0.060	2.845**	0.006	2.055	-0.058	1.274	-0.094*	2.873*	-0.086
	(1.379)	(0.076)	(1.125)	(0.051)	(1.319)	(0.054)	(1.571)	(0.056)	(1.561)	(0.052)
2. Test Incentive	3.826***	-0.147**	2.645**	-0.024	3.378***	-0.080	4.029***	0.070	2.275*	-0.015
	(1.345)	(0.061)	(1.051)	(0.049)	(1.168)	(0.052)	(1.253)	(0.053)	(1.313)	(0.048)
3. (Anemia)X(Test)	-2.665	0.044	-2.855*	0.051	-2.453	0.119	-1.853	0.076	-4.034*	0.002
	(1.943)	(0.091)	(1.646)	(0.077)	(1.871)	(0.080)	(2.069)	(0.083)	(2.138)	(0.080)
4. Observations	1328	1338	1277	1284	1272	1273	1067	1076	1133	1134
5. R-squared	0.272	0.576	0.214	0.587	0.248	0.591	0.233	0.584	0.257	0.580
6. p-Value [H_0:Anemia = Test]	0.310	0.140	0.860	0.530	0.250	0.630	0.040**	0.000***	0.680	0.210
7. Dual Incentive	3.860***	-0.160**	2.640**	0.030	2.980**	-0.020	3.450***	0.050	1.110	-0.100**
8. p-Value [H_0:Dual = 0]	0.000	0.020	0.010	0.490	0.010	0.730	0.000	0.340	0.350	0.040
9. Control Group (No Incentive) Mean	128.160	0.070	132.070	-0.040	136.120	-0.090	139.150	-0.020	144.860	0.050
Panel B: Large Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	0.273	-0.038	-0.461	-0.028	-0.768	0.033	-3.311***	-0.127**	-3.115***	-0.036
	(1.217)	(0.055)	(1.218)	(0.054)	(1.376)	(0.049)	(1.208)	(0.058)	(1.041)	(0.048)
2. Test Incentive	-0.875	0.023	-0.848	-0.133**	-2.242	0.028	-2.349*	-0.001	1.016	0.121**
	(1.331)	(0.064)	(1.205)	(0.051)	(1.446)	(0.047)	(1.388)	(0.064)	(1.188)	(0.058)
3. (Anemia)X(Test)	1.719	0.041	1.454	0.102	2.644	0.003	3.209*	0.070	1.918	-0.154**
	(1.588)	(0.075)	(1.503)	(0.071)	(1.811)	(0.063)	(1.643)	(0.081)	(1.597)	(0.070)
4. Observations	1145	1144	1243	1242	1296	1291	1146	1144	1203	1198
5. R-squared	0.293	0.645	0.155	0.614	0.211	0.611	0.222	0.631	0.213	0.590
6. p-Value [H_0:Anemia = Test]	0.440	0.360	0.740	0.030**	0.270	0.920	0.470	0.060*	0.000***	0.020**
7. Dual Incentive	1.120	0.030	0.150	-0.060	-0.370	0.060	-2.450*	-0.060	-0.180	-0.070
8. p-Value [H_0:Dual = 0]	0.280	0.660	0.900	0.350	0.800	0.240	0.080	0.300	0.870	0.160
9. Control Group (No Incentive) Mean	128.740	0.210	133.750	0.090	138.460	-0.090	142.050	-0.010	146.370	0.060

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline exam scores, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.18: Effects of Anemia and Test Incentives by Baseline Hb Quintile

portion of teacher attention. This could explain a less negative effect among students at the lower end of the initial distribution in anemia incentive schools since their health improved more relative to their classmates.

Table 6.19 presents incentive impacts by baseline exam score quintile. Again focusing on small block grant schools, three findings are of note. First, the effects of anemia incentives, test incentives and dual incentives on hemoglobin are relatively constant (and statistically similar) across the quintiles (panel A, Rows 1 & 2). Second, the negative effects of incentives on exam scores are largest for children in the right tail of the initial test score distribution (Columns 8 & 10). A plausible explanation for this finding is that this reflects ceiling effect where the marginal return to additional inputs (including improved iron status) for these children is low.

6.3.3 Mechanisms

6.3.3.1 Use of Block Grants

Results for the planned and reported use of block grants are shown in Table 6.20 (for the pooled sample), Table 6.21 (for small block grant schools) and Table 6.22 (for large block grant schools). Planned block grant use amounts (Panel A) are from a non-binding budget filled by administrators immediately after they learned of incentives and the grants they would receive. Reported block grant use (Panel B) is from an identical budget from filled by administrators at the time of the endline survey. In this table, the reported expenditure per student on each of four categories is regressed on treatment dummies and interactions as in Equation (5.4). Due to limited power for regressions at the school level, however, these regressions do not include county and randomization strata dummies. [14]

I examine two nutrition related spending categories (foodstuff, supplements/fortified flour and soy), education-related expenditures, and one category that could be either (communications and transfers to households). "Foodstuff" includes all school purchases of food (regardless of iron content). "Supplements/fortified flour and soy" includes school spending on supplements

[14] This is a deviation from the pre-analysis plan. Including controls as pre-specified subsumes degrees of freedom and reduces the efficiency of school-level regressions.

	Q1		Q2		Q3		Q4		Q5	
	Hb	Test Score	Hb	Test Score	Hb	Test Score	Hb	Test Score	Hb	Test Score
Panel A: Small Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	2.514*	-0.052	2.296	-0.024	3.022***	-0.024	1.690	-0.006	2.647**	-0.071*
	(1.428)	(0.056)	(1.454)	(0.058)	(1.069)	(0.039)	(1.172)	(0.041)	(1.146)	(0.038)
2. Test Incentive	2.756**	0.087	3.854***	0.025	4.057***	0.019	3.646***	-0.110***	3.584***	-0.110**
	(1.261)	(0.055)	(1.092)	(0.052)	(0.984)	(0.049)	(1.143)	(0.041)	(1.137)	(0.043)
3. (Anemia)X(Test)	-4.035*	0.046	-2.291	0.042	-3.326**	-0.019	-2.058	0.066	-3.278**	0.106*
	(2.216)	(0.081)	(1.946)	(0.077)	(1.387)	(0.065)	(1.525)	(0.062)	(1.535)	(0.055)
4. Observations	1170	1669	1165	1638	1309	1857	1268	1806	1164	1688
5. R-squared	0.378	0.195	0.408	0.144	0.413	0.137	0.421	0.114	0.447	0.165
6. p-Value [H_0:Anemia = Test]	0.880	0.010**	0.210	0.370	0.340	0.310	0.140	0.010**	0.440	0.160
7. Dual Incentive	1.230	0.080	3.860***	0.040	3.750***	-0.020	3.280***	-0.050	2.950***	-0.070*
8. p-Value [H_0:Dual = 0]	0.340	0.110	0.000	0.390	0.000	0.540	0.010	0.270	0.000	0.060
9. Control Group (No Incentive) Mean	136.480	-1.120	135.990	-0.440	135.450	0.050	137.430	0.480	136.320	1.010
Panel B: Large Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	-2.928**	-0.119**	-2.335*	0.032	-0.750	0.026	-1.324	-0.089*	-0.890	-0.117***
	(1.405)	(0.053)	(1.262)	(0.043)	(1.124)	(0.040)	(1.289)	(0.050)	(1.265)	(0.038)
2. Test Incentive	-2.048	0.018	-3.013**	-0.076	-2.297**	0.014	-0.227	-0.034	-1.198	-0.023
	(1.416)	(0.056)	(1.358)	(0.057)	(1.148)	(0.043)	(1.495)	(0.042)	(1.376)	(0.038)
3. (Anemia)X(Test)	4.351**	0.035	4.140**	0.038	0.255	-0.016	2.075	0.057	2.294	0.086
	(1.848)	(0.072)	(1.653)	(0.062)	(1.460)	(0.066)	(1.696)	(0.060)	(1.855)	(0.054)
4. Observations	1185	1697	1225	1743	1284	1771	1174	1644	1164	1579
5. R-squared	0.354	0.174	0.320	0.137	0.360	0.126	0.372	0.124	0.391	0.197
6. p-Value [H_0:Anemia = Test]	0.550	0.010**	0.580	0.050*	0.200	0.810	0.470	0.190	0.820	0.010**
7. Dual Incentive	-0.620	-0.070	-1.210	-0.010	-2.790**	0.020	0.520	-0.070	0.210	-0.050
8. p-Value [H_0:Dual = 0]	0.670	0.170	0.320	0.890	0.020	0.650	0.680	0.210	0.860	0.140
9. Control Group (No Incentive) Mean	139.240	-1.090	139.060	-0.490	138.290	0.080	137.920	0.520	136.060	1.040

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline exam scores, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.19: Effects of Anemia and Test Incentives by Baseline Exam Score Quintiles

	Mean in Control (No Incentive) Group	Coefficient (standard error) on:			N
		Anemia Incentive	Test Incentive	(Anemia)X (Test)	
A. Planned Block Grant Use (Per Student)					
1. Foodstuff	21.518	-2.426 (4.153)	8.045 (10.061)	-6.874 (11.079)	260
2. Supplements/Fortified Flour, Soy	15.192	3.280 (2.345)	1.447 (2.465)	0.823 (3.565)	260
3. Communication/Transfers to Households	7.159	0.449 (2.115)	-1.512 (1.594)	-1.151 (2.419)	260
4. Education	16.497	-1.817 (3.149)	0.631 (3.046)	-1.934 (4.151)	260
B. Reported Block Grant Use at Endline (Per Student)					
5. Foodstuff	16.380	-3.603 (3.317)	-0.669 (4.839)	-1.626 (5.672)	260
6. Supplements/Fortified Flour, Soy	19.011	0.269 (2.719)	0.605 (2.862)	3.121 (4.044)	260
7. Communication/Transfers to Households	4.337	1.627 (1.418)	-0.297 (0.956)	-2.259 (1.663)	260
8. Education	13.345	5.084 (3.626)	3.753 (3.722)	-10.366** (4.939)	260

NOTES. Data source: authors' survey. Table shows school administrator "planned" (Panel A) and "reported" (Panel B) use of block grant funds recorded by having administrators fill out a (non-binding) budget at the onset of the nutrition program and the same budget form reporting actual use of funds at endline. Each row shows coefficient estimates (and robust standard errors) from a seperate regressions estimated using equation (5.4) but excluding county and randomization strata dummies. Dependent variables, the reported expenditure of grants by category (per student), are shown at left. "Foodstuff" includes spending on any type of food purchased using block grant funds; "Vitamins/Fortified Flour, Soy" included purchases of multivitamins or fortified flour or soy sauce; "Communication/Transfers to Households" includes funds used for parent meetings, informational materials for parents, or direct transfers to households (in cash or kind); "Education" includes spending on education expenditures teacher training, teacher overtime, tutoring fees for students, academic supplies (computers, stationary, desk chairs, etc) and incentives/prizes for teachers and students. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.20: Planned and Reported Use of Block Grants - Full Sample

containing iron and fortified flour and soy. Higher spending in this category may suggest an approach of school administrators that was more narrowly focused on iron delivery. "Communications/Transfers to households" includes spending on parent meetings (possibly related to nutrition, but, also possibly not related to nutrition), other communication or nutrition related education materials and transfers such as food subsidies given directly to households (direct transfers were rare). The "Education" category includes spending on teacher training, teacher overtime, tutoring fees, academic supplies (computers, stationary, chairs, desks, etc), and other spending identified by administrators as education-related.

Though there is limited power to detect school level differences, several patterns can be seen in these estimates. First administrators in anemia incentive schools appear to have planned to spend less on educational inputs compared to administrators with test incentives. These differences,

	Mean in Control (No Incentive) Group	Coefficient (standard error) on:			N
		Anemia Incentive	Test Incentive	(Anemia)X (Test)	
A. Planned Block Grant Use (Per Student)					
1. Foodstuff	8.410	2.869 (3.366)	1.575 (2.794)	-4.272 (4.190)	130
2. Supplements/Fortified Flour, Soy	13.040	3.434 (3.276)	-2.529 (2.962)	1.611 (4.111)	130
3. Communication/Transfers to Households	4.753	-2.139 (1.653)	1.262 (1.972)	-0.413 (2.235)	130
4. Education	9.793	-4.189* (2.482)	-0.404 (2.586)	2.846 (3.297)	130
B. Reported Block Grant Use at Endline (Per Student)					
5. Foodstuff	6.262	-1.389 (2.422)	1.704 (2.683)	-2.878 (3.266)	130
6. Supplements/Fortified Flour, Soy	17.074	1.202 (3.243)	-3.850 (3.186)	3.000 (4.333)	130
7. Communication/Transfers to Households	2.462	1.126 (1.192)	0.944 (1.000)	-2.439 (1.544)	130
8. Education	7.513	0.691 (2.535)	-0.883 (2.427)	0.773 (3.517)	130

NOTES. Data source: authors' survey. Table shows school administrator "planned" (Panel A) and "reported" (Panel B) use of block grant funds recorded by having administrators fill out a (non-binding) budget at the onset of the nutrition program and the same budget form reporting actual use of funds at endline. Each row shows coefficient estimates (and robust standard errors) from a seperate regressions estimated using equation (5.4) but excluding county and randomization strata dummies. Dependent variables, the reported expenditure of grants by category (per student), are shown at left. "Foodstuff" includes spending on any type of food purchased using block grant funds; "Vitamins/Fortified Flour, Soy" included purchases of multivitamins or fortified flour or soy sauce; "Communication/Transfers to Households" includes funds used for parent meetings, informational materials for parents, or direct transfers to households (in cash or kind); "Education" includes spending on education expenditures teacher training, teacher overtime, tutoring fees for students, academic supplies (computers, stationary, desk chairs, etc) and incentives/prizes for teachers and students. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.21: Planned and Reported Use of Block Grants - Small Grant Schools

however, are not as clear in the reported use of block grants. Second, administrators with test incentives planned to spend less on more focused iron delivery (supplements and iron fortified flour and soy) and more on food compared to anemia incentive. Administrators with test incentives initially allocated somewhat less to supplements in small block grant schools and more to food in large block grant schools. Higher allocation to food in large block grant schools may be due to a belief among administrators that broader nutritional gains may improve academic performance more than addressing iron deficiency alone. Finally, administrators with dual incentives in large block grant schools report spending significantly less of the grant on education expenses at endline.

	Mean in Control (No Incentive) Group	Coefficient (standard error) on: Anemia Incentive	Test Incentive	(Anemia)X (Test)	N
A. Planned Block Grant Use (Per Student)					
1. Foodstuff	34.228	-6.755 (6.620)	13.500 (17.342)	-9.359 (19.347)	130
2. Supplements/Fortified Flour, Soy	17.280	2.221 (3.369)	4.038 (3.728)	1.680 (5.662)	130
3. Communication/Transfers to Households	9.493	3.609 (3.791)	-3.581 (2.568)	-2.773 (4.265)	130
4. Education	22.999	1.004 (4.849)	1.680 (4.765)	-6.154 (6.623)	130
B. Reported Block Grant Use at Endline (Per Student)					
5. Foodstuff	26.191	-4.750 (5.118)	-2.110 (8.007)	-1.023 (9.502)	130
6. Supplements/Fortified Flour, Soy	20.889	-1.836 (4.452)	3.703 (4.458)	5.173 (6.763)	130
7. Communication/Transfers to Households	6.155	2.533 (2.649)	-1.199 (1.637)	-2.721 (3.135)	130
8. Education	19.001	9.237 (6.014)	7.131 (6.281)	-19.924** (8.088)	130

NOTES. Data source: authors' survey. Table shows school administrator "planned" (Panel A) and "reported" (Panel B) use of block grant funds recorded by having administrators fill out a (non-binding) budget at the onset of the nutrition program and the same budget form reporting actual use of funds at endline. Each row shows coefficient estimates (and robust standard errors) from a seperate regressions estimated using equation (5.4) but excluding county and randomization strata dummies. Dependent variables, the reported expenditure of grants by category (per student), are shown at left. "Foodstuff" includes spending on any type of food purchased using block grant funds; "Vitamins/Fortified Flour, Soy" included purchases of multivitamins or fortified flour or soy sauce; "Communication/Transfers to Households" includes funds used for parent meetings, informational materials for parents, or direct transfers to households (in cash or kind); "Education" includes spending on education expenditures teacher training, teacher overtime, tutoring fees for students, academic supplies (computers, stationary, desk chairs, etc) and incentives/prizes for teachers and students. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.22: Planned and Reported Use of Block Grants - Large Grant Schools

6.3.3.2 Nutrition and Education Inputs

The evidence in the previous section was on how block grants were spent. Tables 6.23 and 6.24 show what student report receiving using data from the student survey.

Table 6.23 shows student-reported receipt of nutrition related inputs - namely, supplements and meat and fruit at school and at home. Again, I show estimates separately for the pooled sample (Panel A), small block grant schools (Panel B) and large block grant schools (Panel C). Two items are of note here. First, among small block grant schools, students in schools where administrators had anemia and test incentives report receiving more of these nutrition related inputs overall compared to no incentive schools based on the summary index (Column 7). Second, consistent with the budget allocations above, students in test incentive schools are significantly less likely to report receiving supplements or meat at home compared to anemia incentive schools (Columns 2 and 5). Instead, the overall gain in inputs in test incentive schools appears to come from relatively more feeding at school. Again, it appears that - even though changes in hemoglobin concentration were similar across anemia and test incentive schools - administrators in these schools used grants to invest more broadly in nutrition instead of more focused iron delivery approaches.

Turning to education inputs in Table 6.24, I find little evidence of any positive effect due to either incentive. In fact, if anything, the point estimates indicate that the anemia incentive had a negative effect on these intermediate outcomes. In the pooled sample and small block grant group estimates, anemia incentives appear to have significantly increased the student reports of teacher absenteeism. These results are consistent with those found for impacts on standardized exam scores and provide further evidence for a reallocation away from education inputs due to the anemia incentive. It is worth noting, however, that there is little evidence of a negative effect due to test incentives.

As mentioned above, incentives could also lead to a reallocation of inputs across children. I test this hypothesis in Tables 6.25 and 6.26. Table 6.25 estimates treatment effects on summary indices for anemia (odd columns) and education (even columns) inputs for each quintile of the baseline hemoglobin distribution. Table 6.26 does the same for each quintile of the baseline

	School provided supplements to children (Child Response)	Times per week supplements distributed by school (Child Response)	Times consumed meat at SCHOOL in past week	Times consumed fruit at SCHOOL in past week	Times consumed meat at HOME in past week	Times consumed fruit at HOME in past week	Summary Index
Panel A: Full Sample (Small and Large Block Grant)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Anemia Incentive	0.022	0.277	0.197	-0.369*	0.233	-0.174	0.041
	(0.045)	(0.215)	(0.223)	(0.193)	(0.202)	(0.272)	(0.033)
2. Test Incentive	0.041	0.204	0.301*	-0.004	0.021	-0.369	0.059*
	(0.043)	(0.199)	(0.166)	(0.212)	(0.196)	(0.280)	(0.032)
3. (Anemia)X(Test)	0.008	-0.065	-0.176	0.474*	0.105	0.683*	-0.006
	(0.055)	(0.285)	(0.235)	(0.270)	(0.270)	(0.383)	(0.043)
4. Observations	16865	17011	18120	18120	18120	18120	18120
5. R-squared	0.340	0.314	0.304	0.181	0.111	0.104	0.167
6. p-Value [H_0:Anemia = Test]	0.640	0.710	0.530	0.060*	0.280	0.480	0.570
7. Dual Incentive	0.070	0.420**	0.320	0.100	0.360**	0.140	0.090***
8. p-Value [H_0:Dual = 0]	0.110	0.040	0.110	0.640	0.060	0.620	0.000
9. Control Group (No Incentive) Mean	0.870	2.710	1.040	1.250	4.090	7.550	-0.030
Panel B: Small Block Grant							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Anemia Incentive	0.172**	0.749**	0.333*	-0.005	0.476**	-0.020	0.113**
	(0.069)	(0.302)	(0.193)	(0.209)	(0.214)	(0.341)	(0.050)
2. Test Incentive	0.110	0.193	0.476***	0.241	0.009	-0.498	0.121***
	(0.077)	(0.311)	(0.158)	(0.230)	(0.230)	(0.363)	(0.046)
3. (Anemia)X(Test)	-0.068	-0.421	-0.409*	0.272	0.073	0.642	-0.074
	(0.092)	(0.413)	(0.236)	(0.299)	(0.335)	(0.533)	(0.064)
4. Observations	8557	8638	9132	9132	9132	9132	9132
5. R-squared	0.395	0.432	0.390	0.188	0.122	0.117	0.217
6. p-Value [H_0:Anemia = Test]	0.260	0.040**	0.430	0.150	0.040**	0.180	0.870
7. Dual Incentive	0.210***	0.520	0.400**	0.510**	0.560**	0.120	0.160***
8. p-Value [H_0:Dual = 0]	0.000	0.100	0.010	0.030	0.010	0.740	0.000
9. Control Group (No Incentive) Mean	0.800	2.560	0.850	1.000	3.880	7.280	-0.110
Panel C: Large Block Grant							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Anemia Incentive	-0.085*	-0.155	-0.198	-0.565**	0.080	-0.613*	-0.037
	(0.050)	(0.285)	(0.277)	(0.240)	(0.327)	(0.342)	(0.035)
2. Test Incentive	-0.078	-0.257	0.046	-0.056	0.117	-0.167	0.000
	(0.047)	(0.229)	(0.239)	(0.283)	(0.303)	(0.399)	(0.039)
3. (Anemia)X(Test)	0.123*	0.466	0.275	0.292	-0.018	0.713	0.030
	(0.071)	(0.381)	(0.292)	(0.337)	(0.397)	(0.505)	(0.050)
4. Observations	8308	8373	8988	8988	8988	8988	8988
5. R-squared	0.557	0.453	0.294	0.248	0.125	0.112	0.257
6. p-Value [H_0:Anemia = Test]	0.890	0.700	0.290	0.080*	0.900	0.250	0.310
7. Dual Incentive	-0.040	0.050	0.120	-0.330	0.180	-0.070	-0.010
8. p-Value [H_0:Dual = 0]	0.320	0.810	0.700	0.260	0.570	0.870	0.860
9. Control Group (No Incentive) Mean	0.940	2.870	1.240	1.500	4.300	7.820	0.050

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.23: Effects of Anemia and Test Incentives on Nutrition Inputs

	Student Absent Last Week (0/1)	Teacher Absences Last Month	Times Student Distracted During Class (last week)	Number HW Assigned Last Week	Number HW Corrected by Teacher	Times Asked Question by Teacher in Class Last Week	Summary Index
Panel A: Full Sample (Small and Large Block Grant)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Anemia Incentive	-0.011	0.574**	0.021	-0.075	-0.260	-0.370*	-0.053*
	(0.012)	(0.260)	(0.045)	(0.215)	(0.252)	(0.210)	(0.027)
2. Test Incentive	-0.008	0.155	-0.057	0.124	0.286	0.238	0.020
	(0.013)	(0.231)	(0.045)	(0.217)	(0.258)	(0.215)	(0.027)
3. (Anemia)X(Test)	0.015	-0.613*	0.067	-0.185	-0.183	0.115	0.015
	(0.017)	(0.348)	(0.068)	(0.311)	(0.347)	(0.296)	(0.040)
4. Observations	18096	17002	17000	16816	16974	16797	18120
5. R-squared	0.047	0.083	0.127	0.228	0.197	0.208	0.159
6. p-Value [H_0:Anemia = Test]	0.820	0.090*	0.100	0.360	0.020**	0.010**	0.010**
7. Dual Incentive	0.000	0.120	0.030	-0.140	-0.160	-0.020	-0.020
8. p-Value [H_0:Dual = 0]	0.730	0.620	0.540	0.510	0.580	0.940	0.520
9. Control Group (No Incentive) Mean	0.180	1.600	0.680	8.270	7.190	5.060	0.020
Panel B: Small Block Grant							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Anemia Incentive	0.003	0.876**	-0.072	-0.014	-0.180	-0.242	-0.059
	(0.017)	(0.436)	(0.068)	(0.331)	(0.375)	(0.293)	(0.037)
2. Test Incentive	-0.008	0.156	-0.112	0.339	0.421	0.636**	0.051
	(0.014)	(0.317)	(0.081)	(0.351)	(0.386)	(0.288)	(0.035)
3. (Anemia)X(Test)	0.005	-0.549	0.291**	-0.242	-0.465	-0.281	-0.037
	(0.025)	(0.483)	(0.118)	(0.439)	(0.492)	(0.382)	(0.050)
4. Observations	9119	8612	8615	8509	8588	8499	9132
5. R-squared	0.042	0.127	0.156	0.272	0.237	0.210	0.199
6. p-Value [H_0:Anemia = Test]	0.490	0.100	0.610	0.240	0.050*	0.000***	0.000***
7. Dual Incentive	0.000	0.480	0.110	0.080	-0.220	0.110	-0.040
8. p-Value [H_0:Dual = 0]	0.980	0.130	0.130	0.800	0.590	0.700	0.170
9. Control Group (No Incentive) Mean	0.150	1.490	0.710	7.910	6.590	4.480	0.000
Panel C: Large Block Grant							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1. Anemia Incentive	-0.038**	0.342	0.034	0.223	-0.017	-0.050	-0.004
	(0.016)	(0.355)	(0.057)	(0.285)	(0.335)	(0.247)	(0.039)
2. Test Incentive	-0.023	-0.037	-0.058	-0.036	0.575*	0.484*	0.042
	(0.016)	(0.370)	(0.056)	(0.255)	(0.298)	(0.273)	(0.039)
3. (Anemia)X(Test)	0.028	-0.583	-0.061	-0.187	-0.225	-0.128	0.024
	(0.025)	(0.504)	(0.083)	(0.443)	(0.474)	(0.340)	(0.057)
4. Observations	8977	8390	8385	8307	8386	8298	8988
5. R-squared	0.064	0.123	0.126	0.283	0.215	0.238	0.188
6. p-Value [H_0:Anemia = Test]	0.420	0.270	0.070*	0.330	0.050*	0.060*	0.210
7. Dual Incentive	-0.03*	-0.280	-0.090	0.000	0.330	0.310	0.060
8. p-Value [H_0:Dual = 0]	0.050	0.510	0.170	1.000	0.330	0.280	0.150
9. Control Group (No Incentive) Mean	0.210	1.720	0.650	8.650	7.830	5.680	0.050

NOTES. Data source: authors' survey. Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline hemoglobin concentration, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.24: Effects of Anemia and Test Incentives on Education Inputs

distribution of exam scores.

Table 6.25 shows that students in higher quintiles of the baseline hemoglobin distribution report receiving significantly more nutrition inputs due to anemia incentives in small block grant schools. That students in lower quintiles are comparatively less likely to report receiving significantly more inputs implies that there was little target by administrators based on baseline anemia status. A higher level of inputs in upper quintiles due to anemia incentives could also be due to heterogeneity in how households responded to nutritional information received from schools. The table also shows evidence clear evidence of a multitasking effect: students at upper quintiles in anemia incentive schools report receiving significantly less in terms of educational inputs. There is no significant difference in educational inputs for the for the bottom quintile, however. Lack of a positive effect for this group suggests that the reduction in inputs for upper quintiles reflects an overall redistribution from education to nutritional inputs rather than a redistribution to students initially anemic. That there is not a significant reduction also implies that this overall reduction in inputs was offset (likely by improved health fi healthier students attract more inputs, teacher attantion for example).

Another interesting result in this table is that test incentives increased student-reported receipt of anemia inputs (comparably to the effect of anemia incentives) and also led to increases in education inputs for the upper quintiles. Panel A, Row 6 shows that the effects of anemia incentives and test incentives were significantly different for quintiles 2-5. This increase in inputs did not lead to improved exam performance, however (Table 6.18).

Heterogeneous effects on reported input receipt along the baseline exam score distribution is shown in Table 6.26. The effect of anemia incentives on anemia inputs in small block grant schools is similar across the quintiles (Panel A, Row 1, even columns). This is what one would expect since there is no reason to believe administrators would target anemia inputs in this dimension due to anemia incentives. Consistent with the results for impacts on exam scores, the negative effects of anemia incentives on education inputs are larger in the upper quintiles of the exam score distribution. A final observation worth noting is that the (negative) effect of the anemia incentive on education inputs and the (positive) effect of test incentives on education inputs are statistically different from each other across all quintiles.

	Q1		Q2		Q3		Q4		Q5	
	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs
Panel A: Small Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	0.060	-0.023	0.141	-0.075*	0.176*	-0.110**	0.214**	-0.118**	0.191*	-0.101*
	(0.093)	(0.036)	(0.106)	(0.043)	(0.090)	(0.045)	(0.090)	(0.055)	(0.100)	(0.060)
2. Test Incentive	0.033	0.033	0.173	0.004	0.164*	0.013	0.143*	0.102**	0.317***	0.130***
	(0.128)	(0.037)	(0.120)	(0.042)	(0.085)	(0.038)	(0.079)	(0.046)	(0.105)	(0.048)
3. (Anemia)X(Test)	-0.026	-0.042	-0.090	0.016	-0.205*	0.025	-0.292**	-0.020	-0.341**	-0.018
	(0.153)	(0.053)	(0.152)	(0.055)	(0.122)	(0.061)	(0.130)	(0.079)	(0.162)	(0.082)
4. Observations	1333	1411	1281	1370	1269	1346	1073	1130	1128	1196
5. R-squared	0.385	0.238	0.352	0.260	0.342	0.281	0.325	0.234	0.307	0.239
6. p-Value [H_0:Anemia = Test]	0.790	0.100	0.750	0.070*	0.880	0.010**	0.330	0.000***	0.210	0.000***
7. Dual Incentive	0.070	-0.030	0.220**	-0.060	0.140	-0.070**	0.060	-0.040	0.170	0.010
8. p-Value [H_0:Dual = 0]	0.520	0.490	0.030	0.180	0.110	0.040	0.470	0.380	0.100	0.800
9. Control Group (No Incentive) Mean	-0.020	-0.050	-0.160	-0.020	-0.150	0.040	-0.220	0.010	-0.280	-0.010
Panel B: Large Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	0.149	0.024	0.069	0.038	0.062	0.032	-0.081	-0.053	0.137	0.030
	(0.107)	(0.057)	(0.102)	(0.061)	(0.089)	(0.046)	(0.109)	(0.042)	(0.115)	(0.048)
2. Test Incentive	0.162	0.124**	0.262**	0.007	0.203*	0.009	0.172	0.018	0.315***	0.031
	(0.119)	(0.053)	(0.119)	(0.055)	(0.105)	(0.044)	(0.113)	(0.041)	(0.101)	(0.048)
3. (Anemia)X(Test)	-0.054	-0.013	-0.144	0.079	-0.226*	0.042	-0.063	0.078	-0.221	-0.040
	(0.133)	(0.081)	(0.136)	(0.082)	(0.123)	(0.064)	(0.131)	(0.059)	(0.134)	(0.071)
4. Observations	1144	1224	1241	1317	1288	1391	1140	1205	1195	1284
5. R-squared	0.334	0.278	0.302	0.251	0.264	0.188	0.250	0.237	0.258	0.208
6. p-Value [H_0:Anemia = Test]	0.880	0.080*	0.040**	0.580	0.160	0.610	0.010**	0.130	0.090*	0.990
7. Dual Incentive	0.260	0.130*	0.190	0.120*	0.040	0.080*	0.030	0.040	0.230	0.020
8. p-Value [H_0:Dual = 0]	0.020	0.020	0.100	0.080	0.720	0.090	0.820	0.370	0.120	0.670
9. Control Group (No Incentive) Mean	0.080	-0.050	-0.010	0.050	0.020	0.030	0.050	0.070	-0.120	0.120

NOTES. Data source: authors' survey. Dependent variables are summary indices of inputs constructed using the method described in Anderson (2008). Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline value of the outcome variable, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.25: Effects of Anemia and Test Incentives on Nutrition and Education Inputs by Baseline Hb Quintile

	Q1		Q2		Q3		Q4		Q5	
	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs	Anemia Inputs	Test Score Inputs
Panel A: Small Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	0.178*	-0.043	0.194**	0.023	0.190*	-0.033	0.142	-0.102**	0.180*	-0.076**
	(0.099)	(0.057)	(0.097)	(0.037)	(0.108)	(0.038)	(0.100)	(0.050)	(0.101)	(0.037)
2. Test Incentive	0.138	0.057	0.135	0.137***	0.208*	0.057	0.215**	0.048	0.239**	0.013
	(0.119)	(0.047)	(0.114)	(0.032)	(0.117)	(0.038)	(0.106)	(0.048)	(0.110)	(0.037)
3. (Anemia)X(Test)	-0.093	-0.021	-0.121	-0.157***	-0.243	-0.061	-0.190	-0.019	-0.205	-0.002
	(0.138)	(0.067)	(0.142)	(0.054)	(0.152)	(0.055)	(0.143)	(0.069)	(0.151)	(0.050)
4. Observations	1651	1797	1631	1746	1852	1955	1804	1882	1684	1751
5. R-squared	0.355	0.186	0.340	0.194	0.305	0.220	0.350	0.262	0.367	0.271
6. p-Value [H_0:Anemia = Test]	0.700	0.030**	0.510	0.010**	0.850	0.010**	0.370	0.000***	0.510	0.010**
7. Dual Incentive	0.220**	-0.010	0.210**	0.000	0.150	-0.040	0.170	-0.070*	0.210*	-0.060*
8. p-Value [H_0:Dual = 0]	0.010	0.870	0.030	0.920	0.140	0.330	0.100	0.080	0.050	0.070
9. Control Group (No Incentive) Mean	-0.220	-0.110	-0.230	-0.030	-0.200	0.020	-0.140	0.020	-0.120	0.080
Panel B: Large Block Grant										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Anemia Incentive	0.007	-0.001	0.028	0.022	0.126	0.010	0.078	0.007	0.051	-0.032
	(0.091)	(0.048)	(0.101)	(0.048)	(0.103)	(0.046)	(0.096)	(0.053)	(0.104)	(0.056)
2. Test Incentive	0.136	0.101**	0.175	0.038	0.303***	0.040	0.299***	0.037	0.274***	-0.009
	(0.093)	(0.050)	(0.107)	(0.048)	(0.107)	(0.043)	(0.096)	(0.051)	(0.104)	(0.056)
3. (Anemia)X(Test)	-0.010	-0.078	-0.058	-0.011	-0.240*	-0.044	-0.247**	0.062	-0.055	0.155*
	(0.116)	(0.069)	(0.131)	(0.066)	(0.127)	(0.061)	(0.113)	(0.077)	(0.123)	(0.084)
4. Observations	1175	1850	1223	1881	1284	1869	1174	1720	1162	1667
5. R-squared	0.305	0.149	0.272	0.150	0.267	0.174	0.293	0.200	0.277	0.239
6. p-Value [H_0:Anemia = Test]	0.150	0.040**	0.120	0.740	0.050*	0.460	0.000***	0.530	0.010**	0.660
7. Dual Incentive	0.130	0.020	0.140	0.050	0.190*	0.010	0.130	0.110*	0.270*	0.110*
8. p-Value [H_0:Dual = 0]	0.160	0.660	0.190	0.310	0.090	0.910	0.230	0.050	0.030	0.080
9. Control Group (No Incentive) Mean	-0.080	-0.030	0.030	0.030	-0.050	0.090	0.050	0.080	0.070	0.080

NOTES. Data source: authors' survey. Dependent variables are summary indices of inputs constructed using the method described in Anderson (2008). Coefficients and standard errors (in parentheses) shown for treatment group dummy variables and interactions obtained by estimating equation (5.4) (controling for baseline value of the outcome variable, student age, student grade, student sex, number of students in the school, whether the school has a canteen, student teacher ratio, distance to the furthest village served, percent of boarding students, whether the school has implemented the "Free Lunch" policy, county dummy variables, and dummy variables for randomization strata). A child is considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines, WHO 2001). *, **, and *** indicate significance at 10%, 5% and 1%.

Table 6.26: Effects of Anemia and Test Incentives on Nutrition and Education Inputs by Baseline Exam Score Quintile

Chapter 7: Summary and Conclusion

While performance incentives for front-line workers in public service organizations can be useful in increasing effort provision, incentives for managers have greater capacity to also influence the allocation of organizational resources within a manager's decision-making authority. In this dissertation, I study the role of managerial incentives in service delivery by analyzing a large-scale policy experiment in rural China designed to test how incentives for school administrators and increased resources under their control affect the implementation of a school-based nutrition program targeting iron deficiency anemia.

Through this experiment, I study three main questions. First, I study the impact of administrator incentives tied to student nutrition, rewarding administrators based on school level changes in the rate of anemic students over the school year. I find that incentives for school administrators tied to anemia reductions were effective when administrators had a smaller amount of resources with which to implement the program. Offering administrators incentives that paid 125 yuan ($19.84) per anemia case averted (realized payouts were equivalent to approximately 2 months' salary on average) reduced endline anemia rates among students anemic at baseline by 12.5 percentage points (or 34%). Evidence on mechanisms suggests that, absent more program resources, larger incentives led principals to urge caregivers at home to improve the diets of students. This result suggests that incentives for providers tied to outcomes that are jointly produced with beneficiaries can encourage them to influence beneficiaries in order to improve the contracted outcome.

I do not, however, find that small performance incentives, 10% of the strength, were effective (and were significantly less effective than larger incentives). Previous studies from a variety of contexts have shown small incentives or prices leading to large changes in behavior (Kremer & Miguel 2007; Thornton 2008; Banerjee et al. 2010; Cohen & Dupas 2010; Karlan et al. 2011; Duflo et al. 2011). The findings in these studies suggest that the existence of incentives or a price may be more important than the actual level of incentives. My results, however, are more in line with traditional theory that the strength (or price effect) of incentives is important. A key difference could be that, compared to one-off decisions such as purchasing a good, individuals may be more willing to calculate expected payoffs when rewards are based on sustained actions (such as in an employment setting). Another key difference is that information on anemia was provided to all school administrators and hence the incentives did not raise the salience of anemia relative to the comparison group.

Second, I study the effect of increasing the amount of resources available to school administrators to implement an anemia reduction program and how the responses of managers to performance incentives vary with the amount of resources under their control. I find that—even absent explicit incentives to reduce anemia—increasing the amount of resources earmarked for the program under control of school administrators led to sizable reductions. Increasing the size of the block grant from 0.3 yuan ($0.05) per child per day (approximately 7,452 yuan, or $1,183, total on average) to 0.7 yuan per child per day ($0.11—approximately 17,388 yuan or $2,760 total, on average) reduced anemia rates among those initially anemic by 19 percentage points (or 53%). Further, I find that the impact of additional resources on student nutrition does not appear to be due to an increase in resources alone. Increasing grant amounts also led to a more efficient use of resources and encouraged schools to engage with households to improve student diets at home. This finding lends support to the motivational effects of additional autonomy heavily emphasized in the public administration literature (Rose-Ackerman 1986; Perry & Wise 1990). This may also reflect an 'obligation' effect whereby administrators felt an obligation to implement programs effectively precisely because they received a significant amount of resources. Although increasing program resources had the largest impact on anemia (in terms of point estimates), increasing resources was twice as expensive (per anemia case averted) as adding large performance incentives to the smaller block grant.

Increasing the amount of resources available to school administrators was much less effective when they were also given (large) incentives. In other words, I find that managerial incentives and resources under managerial control are substitutes in this context. I find this substiuttion effect not only in results for student nutrition, but also in nutritional inputs (such as iron-containing supplements, iron-rich and iron absorption enhancing foods and nutritional information to children and households) provided by schools. One explanation for substitution is that increasing resources negated the effect that anemia-based incentives had in motivating administrators to mobilize extra-marginal resources (e.g. engage households). This cannot be the sole explanation, however, given the degree of substitution found: in the presence of a large grant, incentives have a slight negative effect on nutrition outcomes. I therefore explore a number of alternative explanations to the extent possible.

Evidence suggests that this 'crowding-out' effect of incentives is unlikely to be due to the way that incentives and budget size convey information about task difficulty (Benabou & Tirole 2003). It is also unlikely to be due to incentives or resources displacing pure intrinsic motivation (such as a sense of organizational mission—Frey & Jegen 2001; Francois & Vlassopoulos 2008; Gneezy et al. 2011; Kamenica 2012). I find strongest empirical support for a mechanism whereby incentive contracts re-frame the task of implementing the nutrition programs from one that was part of the professional role of administrators to one that was not, in line with the general insights of Benabou and Tirole (2006) and Heyman and Ariely (2004). Shifting the frame of the activity from one that was professional and part of administrators role as members of an organization (the educational bureaucracy) could have two effects. First, it could have negated any 'obligation' that administrators felt to effectively implement programs due to increased resources. With an incentive structures as personal income, administrators could 'pay' to avoid the obligation in foregone rewards. Second, it may have increased (or administrators perceived them as increasing) perceptions of their liability for potentially negative outcomes (the risk of food poisoning, for example). Because smaller budgets, in practice, kept administrators from being able to afford food and other inputs beyond iron supplements (which were perceived as safe and were not procured by principals themselves) incentives were effective in this case through effort to enforce “compliance” with a supplement regimen. With a larger budget and incentive, risk-averse and career-concerned school administrators facing ‘bureaucratic incentives’ were reluctant to devote effort or resources to the nutrition program. This result is in conflict with the general view that incentives and autonomy should be complementary as autonomy provides more flexibility to respond to incentives. More generally, this crowding-out result suggests that incentives could alter how tasks or outcomes are attributed to managers being incentivized. Future research should directly test how incentives are interpreted and the consequences of this in an organizational context.

The third main question I study is how provider (school administrator) incentives to reduce anemia and improve test scores each affect anemia prevalence and academic performance. Although the theory of multitasking (following from the work of Holmstrom and Milgrom 1991) is well developed, there are few empirical studies testing this theory directly. The context of school-based nutrition programs allows me to simultaneously test how incentives for a given ’direction’ of effort affect outcomes that are more or less well-aligned. Because student nutrition is an input into education production but academic performance is not an input into nutritional production, test incentives are well-aligned with improving nutrition but anemia-based incentives are not well aligned with effort to purely increase academic performance. I also test the effect of rewarding multiple outcomes simultaneously and asses the potential for multiple rewards to lessen distortionary effects due to multitasking. I emphasize three main findings. First, incentives in the two dimensions (given in the context of an anemia reduction program) both led to significant reductions in anemia prevalence. Second, anemia-based and test-based incentives serve as substitutes in the direction of anemia reduction: providing administrators with both types of incentives did not lead to significantly larger reductions in anemia. Third, I find that anemia incentives caused an allocation of resources away from education ’inputs’ but this did not lead to significantly lower student performance on standardized exams after one year. Less reallocation occurred when administrators had test based incentives or both types of incentives. Collectively, these results confirm theoretical predictions and reflect that test-based incentives are well-aligned with improving nutrition, but anemia-based incentives are not well aligned with effort to improve academic performance. Strengthening incentives to improve academic performance while also emphasizing the relationship between good nutrition and academic performance may therefore be sufficient to motivate administrators to effectively implement school-based anemia reduction programs while causing less reallocation of resources away from education. The result that rewarding multiple dimensions simultaneously was not additionally beneficial also has implications for the design of performance incentive schemes that use multiple measures in reward determination (as is common in health applications). Rewarding additional measures that are jointly produced may not lead to additional increases in contracted outcomes.

A number of caveats should be considered along with the results presented here. First, there are presumably more optimal contract structures than the ones used. The contracts used were chosen for their simplicity and transparency. Optimal contract design also requires more information than is typically available (information on risk aversion, for example). At the same time, it is unclear if more complex contracts can produce better outcomes. Any potential gains from more optimal contracts may be offset by responses to complexity or lack of transparency. This is an important area for future research.

Second, I only study effects of incentives over the course of one school year. How effects evolve over time, however, could be critical. Although ratchet-type effects are unlikely given a linear contract, the motivational effects of outcome-based contracts could degrade over time due to variability in outcome measures that is uncorrelated with effort. This could happen, for instance, if providers who invest significant effort are not rewarded due to downside shocks in the outcome measure. Similarly, providers who chose not to allocate substantial effort may nevertheless be rewarded substantially due to upside shocks. In either instance, effort could decline in subsequent periods. Future research should investigate this issue directly, including the comparison of rewards tied to less-variable "input" measures and outcome-based incentives over time. While input incentives may not be as closely related to the outcomes ultimately valued and do not create incentives for innovation, this may be compensated by better reflection of provider effort.

Finally, as with any empirical study, my results are only internally valid and may vary in other settings, particularly settings with significantly different incentive systems. Future research should examine managerial incentives and their potential to improve service delivery in other geographical and institutional contexts.

Regardless of these caveats, the results presented here offer insights into the design of incentives in public service organizations. Overall, results suggest that performance incentives for managers may hold promise for improving service delivery in developing countries; however, policy makers designing performance incentive schemes should consider carefully institutional features and existing incentives embedded within public service organizations.

Chapter A: Intervention Materials

A.1 Training Materials

A.1.1 Anemia Information Video Script Translation

Hello! My name is Dr. Wang and today I'm going to talk to you about something called anemia. Did you know that around 40% of children in rural areas of China are anemic? This is a problem for principals like yourself because anemic children do worse in school than healthy children!

So, what is anemia? Anemia is a disease that makes it difficult for the oxygen you breathe in to get to your muscles and brain. Without oxygen in your brain, you are unable to think clearly. So how do we fix this? Well, by eating lots of iron-rich foods! Foods like meat, beans and tofu will help turn a sickly, anemic person into a stronger, smarter healthy person!

How can you tell if any students in your school are anemic?

One of the most troublesome things about anemia is that there are often no symptoms. A healthy looking child may still be anemic! At times, there may be signs that children are anemic. Anemic children sometimes get headaches and complain about being dizzy. Anemic children sometimes feel chronically tired. They might be pale. Anemic students may have a hard time paying attention in class. Many anemic children end up skinner and shorter than they should be. BUT SOME DO NOT. Again: there may be children that do not have any of these symptoms, but who still have anemia. And, when students have anemia, this puts them at a disadvantage compared to their healthier classmates.

In fact, researchers have studied anemia and its effect on grades in Shaanxi, Ningxia, Gansu and Qinghai. In ALL of the studies, it was found that children with anemia perform worse than their classmates in school. They get worse grades, and perform worse on standardized intelligence tests. When anemic students are treated for anemia, their grades in school have been shown to increase by up to five points! This is the difference between a C+ (75 points out of a 100) and a B (80 points or more)!

There is some Good News, however. All of these problems can be easily fixed, just by treating anemia in your school.

Researchers have found that around one-third of children in your area have anemia. If you suspect that some students in your school might be anemic, don't worry—anemia is easy to treat! To fight and prevent anemia—and improve student performance in class—you can just make a few simple changes in your students' diet to make sure they are getting enough iron and eating a balanced diet.

What foods have iron? MEAT. Chicken, beef, lamb and pork—all animal meats are great sources of iron. Eggs and milk are healthy foods, but they don't have any iron!

Some other foods have iron too, but it is harder for children to absorb iron from non-meat sources. Meat is the best source of iron. Other sources of iron are pumpkin or squash seeds, potato skins, tofu and soy products, peanuts, and beans. If you want to add these foods to your students' diet, make sure you also add lots of fruits and vegetables so that the children can absorb the iron into their bodies. Vegetables like green peppers, chili peppers, and red dates are all good sources of vitamin C, which children need to absorb iron. Did you know that adding a red date or other source of vitamin C to your students' meal can boost their rate of iron absorption by as much as 300%?

Eating a balanced diet each day is important for overcoming anemia, so feed your child lots of meat and vegetables in addition to staples like plain rice, noodles, and buns. Eating staples may make children feel full, but they don't contain any iron and have few other nutrients, so by themselves they are not healthy. Staples should be paired with meat or vegetables to make a healthy meal.

Remember, adding ONLY more fruits and vegetables to your students' meals is good, but not enough.

Adding ONLY more grains to your students' meals is good, but not enough.

A balance of meat, fruits and vegetables is key to keeping students in your school healthy and ready to learn!

Exactly how much meat does a child need to eat to stay healthy? He should eat half a jin of meat each day. How much is that? Make a fist with your hand. Every day your students should eat about two fists of meat. More specifically, each student could eat: • $\frac{1}{2}$ jin of pork (two and a half fists) • 1/3 jin of chicken (one and a half fists) • $\frac{3}{4}$ jin of lamb (four fists) • 4/5 jin of tofu (two fists) • 4/3 jin of beans (one bowl)

Keep in mind that this is how much iron a child should consume in a full day, not just in a single meal. Since many children eat meals at both home and school, you may have to share the responsibility for child nutrition with students' parents. You can talk to them about student nutrition and the importance of a balanced meal.

Even if you can't provide students with as much meat as they need, remember that some meat is better than none at all!

Now let's make some healthy meal combinations! Here are some examples of a healthy school lunch:

• Chicken and green pepper • Pork and chopped cabbage • Tofu and chili peppers • String beans with ground pork

In addition to supplementing school meals with meat, vegetables and fruit, there are other things you can do to reduce anemia in your schools. Some of these options might be cheaper or easier than buying meat everyday, and can be just as effective!

One option is to give students a daily multivitamin with iron. A single vitamin tablet has as much iron as a whole dish of pork! Vitamins are easy to dispense – other principals in your area have tried it and found that students in their schools became more energetic and performed better in school. A daily multivitamin helps prevent anemia and especially helps children who are already anemic and most at-risk for learning problems. Follow the recommended dosages for children on the bottle, and be sure to provide clean drinking water for students to use when taking vitamins. Many manufacturers now make chewable vitamins for children. These are easy to take, don't require any water, and kids often enjoy the taste!

Vitamins may not be readily available at your local grocery store, but we can help you purchase them easily.

Another way to reduce anemia in your school is to cook school meals with foods that are fortified with iron. Some special types of soy sauce and flour have extra iron added. These are cheap, and taste the same as regular soy sauce and flour. Fortified foods are often difficult to find at your local grocery store, but we can help you purchase them easily.

You may also choose to tell parents about anemia and why it is important. Tell them that their children might not be doing as well in school as they should be, and share the strategies you have learned to reduce anemia. Explain the importance of a balanced diet. Persistent efforts from teachers and principals to talk to parents is very important, and there are many successful examples of this. Most parents are very concerned about the well-being of their children. However, it is not an easy task to convey key nutrition and health information to parents! Giving parents pamphlets or letters about anemia and how to prevent anemia is not always effective. Some parents may not receive the information, some may not understand it, and some may think they do not have the economic means to change their dietary habits.

Instead, try holding meetings at school with parents to communicate information directly. Frequently follow up to check in about progress and to answer any questions parents might have.

Remember that the best strategy for reducing anemia will depend on local conditions at your school—and no one knows your school better than you do! Can you think of other ways to reduce anemia in your school? Be creative! You know your school and students best, so you may be able to think of even better strategies for keeping your students healthy.

Why is it so important to address the anemia problem in your school?

Well, researchers in Shaanxi, Ningxia, Gansu and Qinghai found that anemic children who took a daily multivitamin with iron saw a big improvement in their test scores. This was because as their anemia was treated, they grew stronger, could think more clearly, and got better grades. Eating meat does the same thing!

Whichever option you choose, remember that it is your responsibility to make sure your students are eating healthily everyday. They will grow taller, think more clearly, and get better grades so that they can go on to high school and college. Help give China's children a better future!

A.1.2 Example Budget Plan

中国科学院地理科学与资源研究所与__________________小学科研合作协议

甲方：中国科学院地理科学与资源研究所；　　　　乙方：__________________小学　　　　甲方将提供人民币______元给乙方用于提高乙方学校学生的营养，健康和教育水平，项目实施时间为2011-2012学年，甲方和乙方将一起商定项目实施细节（见下表），项目经费甲方将分两次拨付给乙方，拨付金额和时间见下表，乙方在收到项目经费后，须参照下表所选内容根据学校实际情况认真做好项目的实施。

校长姓名:　　　　学校代码:　　　　分组代码:

项目预算金额　0　项目可用金额　12600　项目剩余金额　**12600**

	数量	估计单位成本	提供持续时间		提供次数		提供价格	每生提供数量	在校学生数	估计总成本
			学期	周（共24周）	每学期	每周				
减少贫血的可能方案										
学校食物补充										
肉（两）										0
豆腐（两）										0
蔬菜（两）										0
水果（两）										0
豆子（两）										0
其他一（　　）										0
其他二（　　）										0
其他三（　　）										0
维生素（片）							0.3			0
铁强化面粉/酱油										0
减少贫血或提高学生学业表现的可能方案										
家长会（次）										0
给老师的激励										0
给学生的激励										0
提高学生学业表现的可能方案										
教师培训费										0
教师加班费										0
学生辅导费										0
日常教学用品										
课本										0
文具										0
计算机										0
课桌										0
椅子（凳子）										0
其他一（　　）										0
其他二（　　）										0
课外资料										0

补贴力度（元/生/天）	*0.5*
在校学生总数（人）	*210*
补贴天数（天）	*120*
补贴总金额（元）	12600

		发放财物总数	第一次发放		第二次发放	
			时间	数量	时间	数量
维生素	元	0		0		0
	片	0		0		0
现金	元	12600		5250		7350

甲方:中国科学院地理科学与资源研究所　　　　乙方:　　　　小学

甲方负责人签名:　　　　乙方负责人签名:

盖章:　　　　盖章:

时间:　　　　时间:

A.2 Contract Example

中国科学院农业政策研究中心
CENTER FOR CHINESE AGRICULTURAL POLICY
CHINESE ACADEMY OF SCIENCES

中国科学院地理科学与资源研究所与[redacted]小学
关于对该校校长解决该校学生贫血问题和提高该校学生学业表现进行奖励的科研合作协议

甲方：中国科学院地理科学与资源研究所

乙方：[redacted]小学

很高兴通知您贵校已经被选定参与西部贫困地区农村学校克服学生贫血，提高学生学业表现的行动研究项目。作为项目的一部分，甲方会根据乙方校长在成功减少乙方贫血学生人数，以及在成功提高乙方学生语文和数学成绩两个方面的成果对乙方校长的工作进行奖励。奖励的金额将由甲方代表项目捐赠方的意愿在项目实施完成后两周内支付给乙方校长。贫血人数减少和考试成绩增加两个方面奖励实施具体办法如下，最终的奖励金额等于贫血人数减少获得的奖励金额与考试成绩增加获得的奖励之和：

贫血减少奖励办法：在 2011 年 9 月，甲方项目团队在乙方开展了基线调查。在 2011 年 9 月的调查中，甲方项目团队对乙方学生的贫血情况做了检测，检测结果表明乙方有___55___个学生存在贫血。甲方项目团队将在 2012 年春季再次对乙方学生的贫血情况进行检测。如果在 2012 年春季的检测中乙方的贫血学生人数小于 2011 年 9 月检测发现的贫血学生人数，同时您在 2012 年春季仍然担任乙方的校长职务，甲方将根据贫血学生人数的变化给乙方校长提供奖励。乙方每减少十个贫血学生，甲方将给乙方校长 1250 元人民币的奖励（每减少一个贫血学生奖励 125 元）。如果在 2012 年春季的检测中乙方的贫血学生人数大于或等于 2011 年 9 月检测发现的贫血学生人数，甲方将不支付任何奖励给乙方校长。

考试成绩增加奖励办法：在 2011 年 9 月，甲方项目团队在乙方学校开展了基线调查。在 2011 年 9 月的调查中，甲方项目团队对乙方学校学生进行了语文和数学标准化测试（一半学生考语文，一半学生考数学），甲方项目团队的测试结果表明乙方学生的综合平均分为70.13919分。甲方项目团队将在 2012 年春季再次对乙方学生的语文和数学能力进行测试（和 2011 年 9 月考试的难度水平相当）。如果乙方学生综合平均分有所提高，同时您在 2012 年春季仍然担任乙方的校长职务，甲方

1

[redacted]

将根据乙方学生综合平均分的增长幅度给乙方校长提供奖励。乙方学生综合平均分每提高 **1** 分，甲方将给乙方校长 800 元人民币的奖励。如果乙方学生综合平均分没有任何提高，甲方将不支付任何奖励给乙方校长。

如有疑问请通过以下方式垂询：

电话：010-64888990 转 802 或 13520074698（罗仁福），

029-88308337 或 13892833777（史耀疆）

祝您工作顺利！生活愉快！

甲方：中国科学院地理科学与资源研究所　　　乙方：[redacted]小学

[redacted]

Jia 11, Datun Road, Anwai, Beijing 100101, China
Tel: (86)-10-64889440
Home Page: http://www.ccap.org.cn
北京市安定门外大屯路甲11号100101
Fax: (86)-10-64856533
E-mail: ccap@igsnrr.ac.cn

A.3 Contract Translation

Agreement between the Chinese Academy of Sciences Institute for Geographic Sciences and Natural Resources and ____________ Concerning School Administrator Rewards to Reduce Anemia and Raise Academic Performance

Party A: Chinese Academy of Sciences Institute for Geographic Sciences and Natural Resources

Party B: _______________(School Administrator)

We are happy to inform you that you have been selected to participate in a project to prevent anemia and raise academic achievement among students in Western China's impoverished areas. As part of this project, Party A offers Party B a performance bonus based on Party B's ability to reduce the number of anemic students in their school and, independently, to raise students' academic achievement. These rewards are being offered to you by the Center for Chinese Agricultural Policy at the Chinese Academy of Sciences (CCAP), who is managing the project on behalf of project sponsors. This reward will be paid within 2 weeks after the completion of the project. The details of bonus calculations based on reducing the number of anemic students and raising are as follows:

Anemia Reward

In September of 2011, Party A conducted a baseline survey in the school of Party B. As part of this survey, Party A collected information on the anemia status of your students and found that ______ students were anemic. In the Spring of 2012, Party A will return to the school of Party B to once again asses students for anemia. If in Spring 2012 the number of students with anemia has declined compared to the number of students found in September 2011and Party B remains as the school's principal at that time, Party A will provide Party B with a reward. For every 10 student decrease in the number of students with anemia, Party A will pay Party B 1,250 RMB (125 RMB per student). If the number of anemic students increases, Party A will not pay a reward to Party B.

Academic Performance Reward

In September of 2011, Party A conducted a baseline survey in the school of Party B. As part of this survey, Party A administered standardized exams in math and Chinese language based on the national curriculum to students in the school of Party B and found an average combined score of _______. In the Spring of 2012, Party A will return to the school of Party B to once again administer standardized exams in math and Chinese. If in Spring 2012, the average combined score of students has increased compared to the average combined score found in September 2011 and Party B remains as the principal at that time, Party A will provide Party B with a reward. For every 1 point increase in the combined average score of the students in the school of Party B, Party A will pay Party B 800 RMB. If the combined average score decreases, Party A will not pay a reward to Party B.

Please contact the following if you have questions:

Phone XXX-XXXXXXXX (Contact Person:_______)

Email: XXXXXX@XXXX.com.cn

Party A: Chinese Academy of Sciences Institute for Geographic Sciences and Natural Resources

Representative:

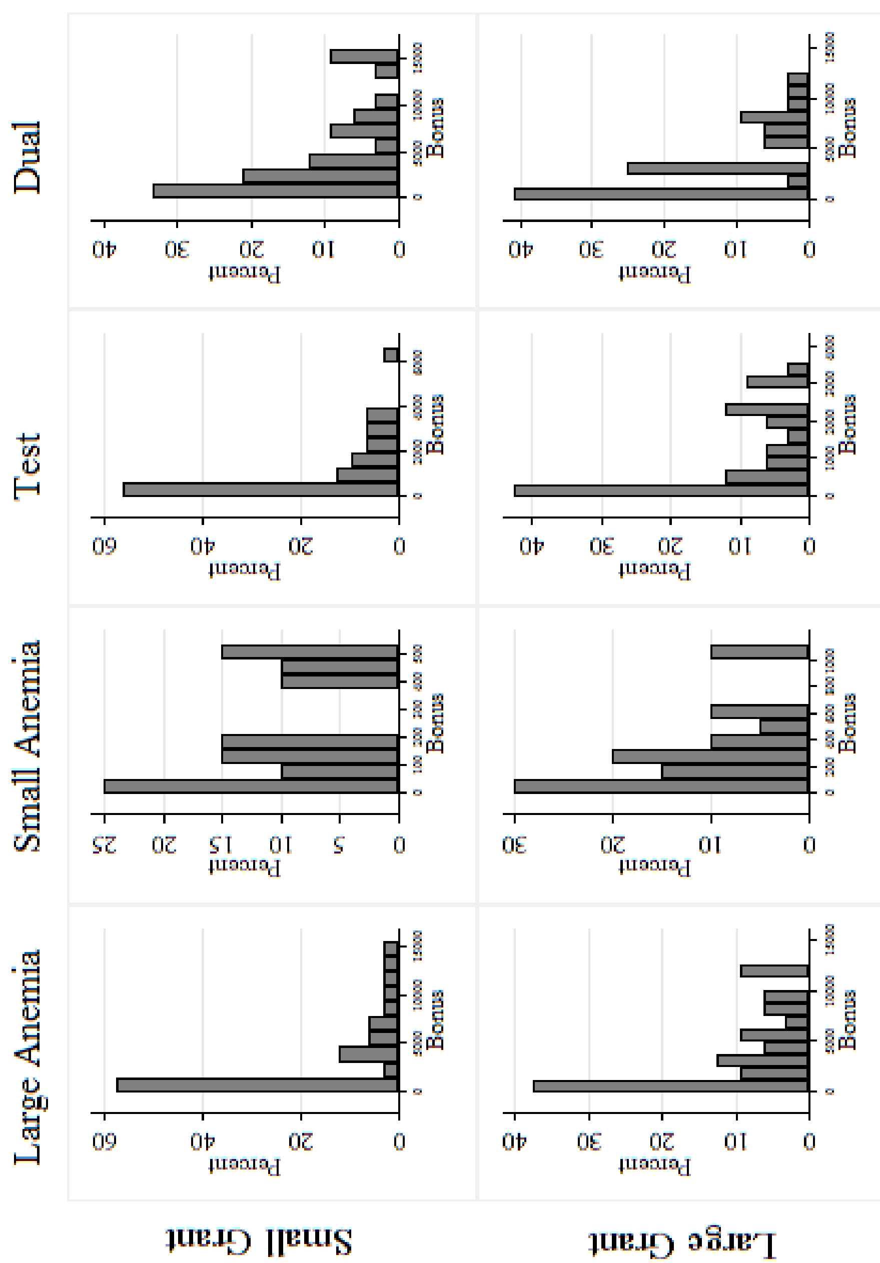
Signed:

Date:

Party B:

Signed:

Date:


Large Anemia
Small Anemia
Test
Dual
Small Grant
Large Grant
Percent
Bonus

Bibliography

Lorena Alcázar, F. Halsey Rogers, Nazmul Chaudhury, Jeffrey Hammer, Michael Kremer, and Karthik Muralidharan. Why are teachers absent? probing service delivery in peruvian primary schools. *International Journal of Educational Research*, 45(3):117–136, 2006.

H. Alderman and D. Bundy. School feeding programs and development: Are we framing the question correctly? *The World Bank Research Observer*, 27(2):204–221, July 2011.

Lindsay H. Allen, Bruno De Benoist, Omar Dary, and Richard Hurrell. Guidelines on food fortification with micronutrients. 2006.

Michael L. Anderson. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495, December 2008.

Anonymous. Education law of the people's republic of china (adopted at the third meeting of the eighth national people's congress),, 1995.

Kenneth J. Arrow. Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5):941–973, December 1963.

Nava Ashraf, Oriana Bandiera, and B. Kelsey Jack. No margin, no mission? a field experiment on incentives for pro-social tasks. *working paper*, 2013.

Orazio Attanasio and Katja Kaufmann. Educational choices, subjective expectations, and credit constraints. *National Bureau of Economic Research Working Paper Series*, No. 15087, July 2009.

George Baker. Distortion and risk in optimal incentive contracts. *The Journal of Human Resources*, 37(4):728–751, October 2002.

George P. Baker. Incentive contracts and performance measurement. *Journal of Political Economy*, 100(3):598–614, June 1992.

Yarlini Balarajan, Usha Ramakrishnan, Emre Özaltin, Anuraj H Shankar, and SV Subramanian. Anaemia in low-income and middle-income countries. *The Lancet*, 378(9809):2123–2135, December 2011.

O. Bandiera, I. Barankay, and I. Rasul. Incentives for managers and inequality among workers: evidence from a firm-level experiment. *The Quarterly Journal of Economics*, 122(2):729–773, 2007.

A. Banerjee and E. Duflo. Addressing absence. *The Journal of Economic Perspectives*, 20(1): 117–132, 2006.

Abhijit Banerjee, Angus Deaton, and Esther Duflo. Health, healthcare, and economic development: Wealth, health, and health survices in rural rajisthan. *The American economic review*, 94(2): 326–330, May 2004.

Abhijit V. Banerjee, Shawn Cole, Esther Duflo, and Leigh Linden. Remedying education: Evidence from two randomized experiments in india. *The Quarterly Journal of Economics*, 122(3):1235–1264, August 2007.

A.V. Banerjee, E. Duflo, R. Glennerster, and D. Kothari. Improving immunisation coverage in rural india: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ: British Medical Journal*, 340, 2010.

World Bank. World development report 2004: Making services work for poor people. Technical report, 2004.

Paulin Basinga, Paul J Gertler, Agnes Binagwaho, Agnes LB Soucat, Jennifer Sturdy, and Christel MJ Vermeersch. Effect on maternal and child health services in rwanda of payment to primary health-care providers for performance: an impact evaluation. *The Lancet*, 377(9775):1421–1428, April 2011.

Jere R. Behrman, Susan W. Parker, Petra E. Todd, and Kenneth I. Wolpin. Aligning learning incentives of students and teachers: Results from a social experiment in mexican high schools. *PIER working paper*, 2013.

Sima Berendes, Peter Heywood, Sandy Oliver, and Paul Garner. Quality of private and public ambulatory health care in low and middle income countries: Systematic review of comparative studies. *PLoS Med*, 8(4):e1000433, April 2011.

T. Besley and M. Ghatak. Reforming public service delivery. *Journal of African Economies*, 16 (1):127, 2007.

E. Bloom, E. King, I. Bhushan, M. Kremer, D. Clingingsmith, B. Loevinsohn, R. Hong, and J. Schwartz. *Contracting for health: evidence from Cambodia*. 2006.

Nicholas Bloom, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. Does management matter? evidence from india. *The Quarterly Journal of Economics*, 128(1):1–51, February 2013.

Miriam Bruhn and David McKenzie. In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1:200–232, September 2009.

D. Bundy and H.L. Guyatt. Schools for health: Focus on health, education and the school-age child. *Parasitology Today*, 12(8):1–14, August 1996.

D.A.P. Bundy, S. Shaeffer, M. Jukes, K. Beegle, A. Gillespie, L. Drake, S.F. Lee, A.M. Hoffman, J. Jones, A. Mitchell, et al. School-based health and nutrition programs. In *Disease Control Priorities in Developing Countries*, volume 30, page 75. World Bank, World Health Organization, National Institutes of Health, New York, 2006.

T.L. Burrows, H. Truby, P.J. Morgan, R. Callister, P.S.W. Davies, and C.E. Collins. A comparison and validation of child versus parent reporting of children's energy intake using food frequency questionnaires versus food records: Who's an accurate reporter? *Clinical Nutrition*, 32(4): 613–618, August 2013.

Margaretha Buurman, Josse Delfgaauw, Robert Dur, and Seth Van den Bossche. Public sector employees: Risk averse and altruistic? *Journal of Economic Behavior & Organization*, 83(3): 279–291, August 2012.

Roland Bénabou and Jean Tirole. Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3):489–520, 2003.

Roland Bénabou and Jean Tirole. Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678, December 2006.

N. Chaudhury, J. Hammer, M. Kremer, K. Muralidharan, and F.H. Rogers. Missing in action: Teacher and health worker absence in developing countries. *The Journal of Economic Perspectives*, 20(1):91–116, 2006.

Nazmul Chaudhury and Jeffrey S. Hammer. Ghost doctors: Absenteeism in rural bangladeshi health facilities. *The World Bank Economic Review*, 18(3):423–441, January 2004.

Jessica Cohen and Pascaline Dupas. Free distribution or cost-sharing? evidence from a randomized malaria prevention experiment. *The Quarterly Journal of Economics*, 125(1):1–45, 2010.

J. Currie, W. Lin, and W. Zhang. Patient knowledge and antibiotic abuse: Evidence from an audit study in china. *Journal of health economics*, 30(5):933–949, 2011.

J. Das, A. Holla, V. Das, M. Mohanan, D. Tabak, and B. Chan. In urban and rural india, a standardized patient study showed low levels of provider training and huge quality gaps. *Health Affairs*, 31(12):2774–2784, 2012.

Jishnu Das and Jeffrey Hammer. Strained mercy: Quality of medical care in delhi. *Economic and Political Weekly*, page 951–961, 2004.

Jishnu Das and Jeffrey Hammer. Money for nothing: The dire straits of medical practice in delhi, india. *Journal of Development Economics*, 83(1):1–36, 2007.

Jishnu Das, Jeffrey Hammer, and Kenneth Leonard. The quality of medical advice in low-income countries. *The Journal of Economic Perspectives*, 22(2):93–114, April 2008.

Jishnu Das, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. School inputs, household substitution, and test scores. *American Economic Journal: Applied Economics*, 5(2):29–57, January 2013.

Amil Dasgupta and Yianis Sarafidis. Managers as administrators: Reputation and incentives. *Journal of Economic Behavior & Organization*, 70(1–2):155–163, May 2009.

Edward L. Deci, Richard Koestner, and Richard M. Ryan. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125 (6):627–668, 1999.

Adeline Delavande, Xavier Giné, and David McKenzie. Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics*, 94(2):151–163, March 2011.

Avinash Dixit. Incentives and organizations in the public sector: An interpretative review. *The Journal of Human Resources*, 37(4):696–727, 2002.

Avinash Dixit. Bureaucracy, its reform and development. *Review of Market Integration*, 4(2): 135–157, August 2012.

Shufa Du, Fengying Zhai, Youfa Wang, and Barry M. Popkin. Current methods for estimating dietary iron bioavailability do not work in china. *The Journal of Nutrition*, 130(2):193–198, January 2000.

Esther Duflo, Michael Kremer, and Jonathan Robinson. Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American Economic Review*, 101(6):2350–2390, October 2011.

Esther Duflo, Rema Hanna, and Stephen P. Ryan. Incentives work: Getting teachers to come to school. *The American Economic Review*, 102(4):1241–1278, 2012.

Pascaline Dupas. Health behavior in developing countries. *Annual Review of Economics*, 3:425–449, September 2011.

Rena Eichler. Can "pay for performance" increase utilization by the poor and improve the quality of health services. *Background papers for the Working Group on Performance Based Incentives. Centre for Global Development*, 2006.

Rena Eichler and Ruth Levine. *Performance Incentives for Global Health: Potential and Pitfalls*. Center for Global Development. Brookings Institution Press: Baltimore, MD, 2009.

Eric R. Eide and Mark H. Showalter. Estimating the relation between health and education: What do we know and what do we need to know? *Economics of Education Review*, 30(5):778–791, October 2011.

Cynthia Eldridge and Natasha Palmer. Performance-based payment: some reflections on the discourse, evidence and unanswered questions. *Health Policy and Planning*, 24(3):160–166, May 2009. PMID: 19202163.

Ernst Fehr and Armin Falk. Psychological foundations of incentives. *European Economic Review*, 46(4–5):687–724, May 2002.

C. Finch. Regulators of iron balance in humans. *Blood*, 84(6):1697–1702, September 1994.

Patrick Francois and Michael Vlassopoulos. Pro-social motivation and the delivery of social services. *CESifo Economic Studies*, 54(1):22–54, March 2008.

Bruno S Frey and Reto Jegen. Motivation crowding theory. *Journal of Economic Surveys*, 15(5): 589–611, December 2001.

Bruno S. Frey, Fabian Homberg, and Margit Osterloh. Organizational control systems and pay-for-performance in the public service. *Organization Studies*, May 2013a.

Bruno S. Frey, Fabian Homberg, and Margit Osterloh. Organizational control systems and pay-for-performance in the public service. *Organization Studies*, May 2013b.

Paul J. Gertler and Christel Vermeersch. Using performance incentives to improve health outcomes. *Working Paper*, June 2012.

Paul Glewwe, Nauman Ilias, and Michael Kremer. Teacher incentives. *American Economic Journal: Applied Economics*, 2(3):205–227, July 2010.

Uri Gneezy and Aldo Rustichini. Pay enough or don't pay at all. *The Quarterly Journal of Economics*, 115(3):791–810, August 2000.

Uri Gneezy, Stephan Meier, and Pedro Rey-Biel. When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives*, 25(4):191–209, 2011.

João Batista Gomes-Neto, Eric A. Hanushek, Raimundo Hélio Leite, and Roberto Cláudio Frota-Bezzera. Health and schooling: Evidence and policy implications for developing countries. *Economics of Education Review*, 16(3):271–282, June 1997.

A.M. Grant. Does intrinsic motivation fuel the prosocial fire? motivational synergy in predicting persistence, performance, and productivity. *Journal of Applied Psychology*, 93(1):48, 2008.

A Hall, E Bobrow, S Brooker, M Jukes, K Nokes, J Lambo, H Guyatt, D Bundy, S Adjei, S T Wen, Satoto, H Subagio, M Z Rafiluddin, T Miguel, S Moulin, J de Graft Johnson, M Mukaka, N Roschnik, M Sacko, A Zacher, B Mahumane, C Kihamia, L Mwanri, S Tatala, N Lwambo, J Siza, L N Khanh, H H Khoi, and N D Toan. Anaemia in schoolchildren in eight countries in africa and asia. *Public Health Nutrition*, 4(3):749–756, June 2001. PMID: 11415481.

Brian J. Hall and Jeffrey B. Liebman. Are CEOs really paid like bureaucrats? *The Quarterly Journal of Economics*, 113(3):653–691, August 1998.

Brian J. Hall and Kevin J. Murphy. The trouble with stock options. *Journal of Economic Perspectives*, 17(3):49–70, 2003.

Eric A. Hanushek and Ludger Woessmann. Overview of the symposium on performance pay for teachers. *Economics of Education Review*, 30(3):391–393, June 2011.

Zahid Hasnain, Nick Manning, and Jan Henryk Pierskalla. The promise of performance pay? reasons for caution in policy prescriptions in the core civil service. *The World Bank Research Observer*, 2014.

James Heyman and Dan Ariely. Effort for payment a tale of two markets. *Psychological Science*, 15(11):787–793, November 2004.

Bengt Holmstrom and Joan Ricart i Costa. Managerial incentives and capital management. *The Quarterly Journal of Economics*, 101(4):835–860, November 1986.

Bengt Holmstrom and Paul Milgrom. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7:24–52, 1991.

Fuhai Hong, Tanjim Hossain, John A. List, and Migiwa Tanaka. Testing the theory of multitasking: Evidence from a natural field experiment in chinese factories. Working Paper 19660, National Bureau of Economic Research, November 2013.

S. Horton and J. Ross. The economics of iron deficiency. *Food Policy*, 28(1):51–75, February 2003.

Hanan G. Jacoby. Is there an intrahousehold 'flypaper effect'? evidence from a school feeding programme. *The Economic Journal*, 112(476):196–221, 2002.

Michael C. Jensen and Kevin J. Murphy. Performance pay and top-management incentives. *Journal of Political Economy*, 98(2):225–264, April 1990.

M. C.H Jukes, L. J Drake, and D. A.P Bundy. *School health, nutrition and education for all: levelling the playing field*. CABI, 2008.

Emir Kamenica. Behavioral economics and psychology of incentives. *Annual Review of Economics*, 4(1):427–452, 2012.

Dean Karlan, John A. List, and Eldar Shafir. Small matches and charitable giving: Evidence from a natural field experiment. *Journal of Public Economics*, 95(5–6):344–350, June 2011.

Eric Konofal, Michel Lecendreux, Juliette Deron, Martine Marchand, Samuele Cortese, Mohammed Zaïm, Marie Christine Mouren, and Isabelle Arnulf. Effects of iron supplementation on attention deficit hyperactivity disorder in children. *Pediatric Neurology*, 38(1):20–26, January 2008.

M. Kremer, N. Chaudhury, F.H. Rogers, K. Muralidharan, and J. Hammer. Teacher absence in india: A snapshot. *Journal of the European Economic Association*, 3(2-3):658–667, 2005.

Michael Kremer and Edward Miguel. The illusion of sustainability. *The Quarterly Journal of Economics*, 122(3):1007–1065, August 2007.

M.R. Kremer, E. Miguel, and R. Thornton. Incentives to learn. *Review of Economics and Statistics*, 91(3):437–456, 2009.

B. Kristjansson, M. Petticrew, B. MacDonald, J. Krasevec, L. Janzen, T. Greenhalgh, GA Wells, J. MacGowan, AP Farmer, B. Shea, et al. School feeding for improving the physical and psychosocial health of disadvantaged students (review). 2009.

A.B. Krueger and D.M. Whitmore. The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project STAR. *The Economic Journal*, 111(468):1–28, 2001.

Jean-Jacques Laffont and Jean Tirole. *A theory of incentives in procurement and regulation*. MIT Press, Cambridge, MA, 1993.

Victor Lavy. Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy*, 110(6):1286–1317, December 2002.

Victor Lavy. Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, 99(5):1979–2011, December 2009.

Edward P. Lazear. *Personnel Economics*. MIT Press, 1995.

Edward P. Lazear. Performance pay and productivity. *The American Economic Review*, 90(5): 1346–1361, December 2000.

Kenneth L Leonard. African traditional healers and outcome-contingent contracts in health care. *Journal of Development Economics*, 71(1):1–22, 2003.

Kenneth L. Leonard and Melkiory C. Masatu. Professionalism and the know-do gap: exploring intrinsic motivation among health workers in tanzania. *Health Economics*, 19(12):1461–1477, 2010a.

Kenneth L. Leonard and Melkiory C. Masatu. Using the hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. *Journal of Development Economics*, 93(2):226–234, November 2010b.

Maureen Lewis. Governance and corruption in public health care systems. *Center for Global Development Working Paper*, 78:1–57, January 2006.

Chengfang Liu, Linxiu Zhang, Renfu Luo, Scott Rozelle, and Prashant Loyalka. The effect of primary school mergers on academic performance of students in rural china. *International Journal of Educational Development*, 30(6):570–585, November 2010.

B. Loevinsohn and A. Harding. Buying results? contracting for health service delivery in developing countries. *The Lancet*, 366(9486):676–681, 2005.

R. Luo, M. Kleiman-Weiner, S. Rozelle, L. Zhang, C. Liu, B. Sharbono, Y. Shi, A. Yue, R. Martorell, and M. Lee. Anemia in rural china's elementary schools: Prevalence and correlates in shaanxi province's poor counties. *Ecology of food and nutrition*, 49(5):357–372, 2010.

R. Luo, X. Wang, L. Zhang, C. Liu, Y. Shi, G. Miller, S. Rozelle, E. Yu, and R. Martorell. Alarmingly high anemia prevalence in western china. *Southeast Asian Journal of Tropical Medicine and Public Health*, 42(5), 2011a.

Renfu Luo, Linxiu Zhang, Chengfang Liu, Qiran Zhao, Yaojiang Shi, Grant Miller, Elaine Yu, Brian Sharbono, Alexis Medina, Scott Rozelle, and Reynaldo Martorell. Anaemia among students of rural china's elementary schools: Prevalence and correlates in ningxia and qinghai's poor counties. *Journal of Health, Population and Nutrition*, 29(5):471–485, 2011b.

Renfu Luo, Yaojiang Shi, Linxiu Zhang, Chengfang Liu, Scott Rozelle, Brian Sharbono, Ai Yue, Qiran Zhao, and Reynaldo Martorell. Nutrition and educational performance in rural china's elementary schools: Results of a randomized control trial in shaanxi province. *Economic Development and Cultural Change*, 60(4):735–772, July 2012a.

Renfu Luo, Yaojiang Shi, Linxiu Zhang, Chengfang Liu, Scott Rozelle, Brian Sharbono, Ai Yue, Qiran Zhao, and Reynaldo Martorell. Nutrition and educational performance in rural china's elementary schools: Results of a randomized control trial in shaanxi province. *Economic Development and Cultural Change*, 60(4):735–772, July 2012b.

Renfu Luo, Yaojiang Shi, Linxiu Zhang, Huiping Zhang, Grant Miller, Alexis Medina, and Scott Rozelle. The limits of health and nutrition education: Evidence from three randomized-controlled trials in rural china. *CESifo Economic Studies*, April 2012c.

Anthea Magarey, Jane Watson, Rebecca K. Golley, Tracy Burrows, Rachel Sutherland, Sarah A. Mcnaughton, Elizabeth Denney-Wilson, Karen Campbell, and Clare Collins. Assessing dietary intake in children and adolescents: Considerations and recommendations for obesity research. *International Journal of Pediatric Obesity*, 6(1):2–11, 2011.

Yukichi Mano, John Akoten, Yutaka Yoshino, and Tetsushi Sonobe. Teaching KAIZEN to small business owners: An experiment in a metalworking cluster in nairobi. *Journal of the Japanese and International Economies*, 2013.

David McKenzie and Christopher Woodruff. What are we learning from business training and entrepreneurship evaluations around the developing world? *The World Bank Research Observer*, 29(1):48–82, 2014.

R.Sue McPherson, Deanna M Hoelscher, Maria Alexander, Kelley S Scanlon, and Mary K Serdula. Dietary assessment methods among school-aged children: Validity and reliability. *Preventive Medicine*, 31(2):S11–S33, August 2000.

Grant Miller and Kimberly Singer Babiarz. Pay-for-performance incentives in low- and middle-income country health programs. *NBER Working Paper*, April 2013.

Grant Miller, Renfu Luo, Linxiu Zhang, Sean Sylvia, Yaojiang Shi, Patricia Foo, Qiran Zhao, Reynaldo Martorell, Alexis Medina, and Scott Rozelle. Effectiveness of provider incentives for anaemia reduction in rural china: a cluster randomised trial. *BMJ : British Medical Journal*, 345, 2012.

Kathleen J. Mullen, Richard G. Frank, and Meredith B. Rosenthal. Can you get what you pay for? pay-for-performance and the quality of healthcare providers. *The RAND Journal of Economics*, 41(1):64–91, March 2010.

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

Kevin J. Murphy. Chapter 38 executive compensation. In *Handbook of Labor Economics*, volume Volume 3, Part B, pages 2485–2563. Elsevier, 1999.

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

P. Nestel. Adjusting hemoglobin values in program surveys. *Washington, DC: International Nutritional Anaemia Consultative Group, ILSI Human Nutrition Institute*, page 2–4, 2002.

C. Nokes, C. van den Bosch, and D.A.P. Bundy. The effects of iron deficiency and anemia on mental and motor performance, educational achievement, and behavior in children. *A report of the INACG. Washington, DC: International Life Sciences Institute*, 1998.

Ministry of Education. The guideline of implementing the nutrition improvement program for compulsory education students in rural china. Technical report, 2012.

Ben Olken, Junko Onishi, and Susan Wong. Should aid reward performance? evience from a field experiment on health and education in indonesia. August 2011.

P. Orazem, P. Glewwe, and H. Patrinos. The benefits and costs of alternative strategies to improve educational outcomes. *Copenhagen Consensus 2008 Challeng Paper: Education*, 2008.

World Health Organization. *Iron Deficiency Anemia: Assessment, Prevention, and Control - A Guide for Program Managers*. World Health Organization, Geneva, Switzerland, 2001.

A.D. Oxman and A. Fretheim. *An overview of research on the effects of results-based financing*. Nasjonalt kunnskapssenter for helsetjenesten (Norwegian Knowledge Centre for the Health Services), 2008.

Paul Oyer and Scott Schaefer. Why do some firms give stock options to all employees?: An empirical examination of alternative theories. *Journal of Financial Economics*, 76(1):99–133, April 2005.

Sant-Rayn Pasricha, Hal Drakesmith, James Black, David Hipgrave, and Beverley-Ann Biggs. Control of iron deficiency anemia in low- and middle-income countries. *Blood*, 121(14):2607–2617, April 2013.

James L. Perry and Lois Recascino Wise. The motivational bases of public service. *Public administration review*, 50(3):367–373, 1990.

Imran Rasul and Daniel Rogger. Management of bureaucrats and public service delivery: Evidence from teh nigerian civil service. *Working Paper*, 2013.

Heather Rose. Do gains in test scores explain labor market outcomes? *Economics of Education Review*, 25(4):430–446, August 2006.

Susan Rose-Ackerman. Reforming public bureaucracy through economic incentives? *Journal of Law, Economics, & Organization*, 2(1):131–161, April 1986.

HPS Sachdev, T. Gera, P. Nestel, et al. Effect of iron supplementation on mental and motor development in children: systematic review of randomised controlled trials. *Public Health Nutrition*, 8(2):117–132, 2005.

Bernard Salanié. The economics of contracts: A primer. *MIT Press Books*, 2005.

Tisamarie B. Sherry, Sebastian Bauhoff, and Manoj Mohanan. Paying for performance in health care: Results from the randomized roll-out of rwanda's national program. Working paper, 2013.

Prakarsh Singh. Performance pay and information: Reducing child malnutrition in urban slums. MPRA Working Paper 29403, University Library of Munich, Germany, 2011.

A. G. Soemantri. Preliminary findings on iron supplementation and learning achievement of rural indonesian children. *The American Journal of Clinical Nutrition*, 50(3):698–702, September 1989.

A. G. Soemantri, E. Pollitt, and I. Kim. Iron deficiency anemia and educational achievement. *The American Journal of Clinical Nutrition*, 42(6):1221–1228, December 1985.

R J Stoltzfus, J D Kvalsvig, H M Chwaya, A Montresor, M Albonico, J M Tielsch, L Savioli, and E Pollitt. Effects of iron supplementation and anthelmintic treatment on motor and language development of preschool children in zanzibar: double blind, placebo controlled study. *BMJ (Clinical Research Ed.)*, 323(7326):1389–1393, December 2001.

Rassamee Sungthong, Ladda Mo-suwan, and Virasakdi Chongsuvivatwong. Effects of haemoglobin and serum ferritin on cognitive function in school children. *Asia Pacific Journal of Clinical Nutrition*, 11(2):117–122, 2002.

Aya Suzuki, Hoang Nam Vu, and Tetsushi Sonobe. Willingness to pay for managerial training: A case from the knitwear industry in northern vietnam. *Journal of Comparative Economics*, July 2013.

Sean Sylvia, Renfu Luo, Linxiu Zhang, Yaojiang Shi, Alexis Medina, and Scott Rozelle. Do you get what you pay for with school-based health programs? evidence from a child nutrition experiment in rural china. *Economics of Education Review*, 37:1–12, December 2013a.

Sean Sylvia, Yaojiang Shi, Hao Xue, Xin Tian, Huan Wang, Qingmei Liu, Alexis Medina, and Scott Rozelle. Survey using incognito standardized patients shows poor quality care in china's rural clinics. *Working Paper*, 2013b.

Justin Tan. Innovation and risk-taking in a transitional economy: A comparative study of chinese managers and entrepreneurs. *Journal of Business Venturing*, 16(4):359–376, July 2001.

Howard Taras. Nutrition and student performance at school. *Journal of School Health*, 75(6): 199–213, August 2005.

D. Thomas, E. Frankenberg, J. Friedman, J. P. Habicht, M. Hakimi, N. Ingwersen, N. Jones, C. McKelvey, G. Pelto, B. Sikoki, et al. Causal effect of health on labor market outcomes: Experimental evidence. *Working Paper*, 2006.

R. L. Thornton. The demand for, and impact of, learning HIV status. *American Economic Review*, 98(5):1829–1863, 2008.

John H. Tyler, Richard J. Murnane, and John B. Willett. Estimating the labor market signaling value of the GED. *The Quarterly Journal of Economics*, 115(2):431–468, 2000.

Ricardo Uauy, Eva Hertrampf, and Manju Reddy. Iron fortification of foods: Overcoming technical and practical barriers. *The Journal of Nutrition*, 132(4):849S–852S, April 2002.

S. Whiting. The cadre evaluation system at the grass roots: The paradox of party rule. In *Naughton B, Yang D, eds. Holding China together: Diversity and national integration in the post-Deng era*, page 101–119. Cambridge University Press, New York, 2004.

Oliver E. Williamson. *The Mechanisms of Governance*. Oxford University Press, January 1996.

James Q. Wilson. *Bureaucracy*. Basic Books, 1989.

Sophie Witter, Atle Fretheim, Flora L Kessy, and Anne Karin Lindahl. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. Technical report, John Wiley & Sons, Ltd, Chichester, UK, February 2012.

Ludger Woessmann. Cross-country evidence on teacher performance pay. *Economics of Education Review*, 30(3):404–418, June 2011.

L. Q. Xu, S. H. Yu, Z. X. Jiang, J. L. Yang, L. Q. Lai, X. J. Zhang, and C. Q. Zheng. Soil-transmitted helminthiases: nationwide survey in china. *Bulletin of the World Health Organization*, 73(4):507–513, 1995.

Ray Yip. Iron deficiency and anemia. In Richard D. Semba and Martin W. Bloem, editors, *Nutrition and Health in Developing Countries*, pages 327–342. Humana Press, January 2001.

Winnie Yip, Timothy Powell-Jackson, Wen Chen, Min Hu, Eduardo Fe, Mu Hu, Weiyan Jian, Ming Lu, Wei Han, and William C. Hsiao. Capitation combined with pay-for-performance improves antibiotic prescribing practices in rural china. *Health Affairs*, 33(3):502–510, March 2014. PMID: 24572187.

Meng Zhao and Paul Glewwe. What determines basic school attainment in developing countries? evidence from rural china. *Economics of Education Review*, 29(3):451–460, June 2010.

Wenhua Zhao, Kyoko Hasegawa, and Junshi Chen. The use of food-frequency questionnaires for various purposes in china. *Public health nutrition*, 5(6A):829–834, 2002.