NBER WORKING PAPER SERIES

CAN BUREAUCRATS REALLY BE PAID LIKE CEOS? SCHOOL ADMINISTRATOR INCENTIVES FOR ANEMIA REDUCTION IN RURAL CHINA

Renfu Luo
Grant Miller
Scott Rozelle
Sean Sylvia
Marcos Vera-Hernández

Working Paper 21302
http://www.nber.org/papers/w21302

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2015

We are grateful to Dana Andersen, Antonio Cabrales, Syngjoo Choi, Adeline Delavande, Katherine Donato, Will Dow, Eric French, Vivian Hoffmann, Rob Jensen, Victor Lavy, Hamish Low, Aprajit Mahajan, Ian Preston, Carol Propper, Imran Rasul, Pedro Rey- Biel, and Alessandro Tarozzi for helpful comments and discussions. We also thank Alexis Medina at the Rural Education Action Program as well as students from the Center for Experimental Economics in Education at Shaanxi Normal University for project support. Soledad Giardili provided excellent research assistance. We are grateful to the National Institutes of Health (Grant Number: R01 HL106023) as well as the Global Under-development Action Fund the Walter H. Shorenstein Asia-Pacific Research Center at Stanford University for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Can Bureaucrats Really Be Paid Like CEOs? School Administrator Incentives for Anemia Reduction in Rural China
Renfu Luo, Grant Miller, Scott Rozelle, Sean Sylvia, and Marcos Vera-Hernández
NBER Working Paper No. 21302
June 2015, Revised December 2015
JEL No. C93,H40,I12,M52,O15

ABSTRACT

A large literature examines performance pay for managers in the private sector, but little is known about performance pay for managers in public sector bureaucracies. In this paper, we study performance incentives rewarding school administrators for reducing anemia among their students. Randomly assigning 170 schools to three performance incentive levels and two orthogonal sizes of unconditional grants, we analyze performance pay and its complementarity with discretionary resources. We find that both large incentives and larger unconditional grants reduced anemia substantially, but incentives were more cost-effective. Performance incentives led administrators to innovate by working with parents, mitigating potentially offsetting compensatory behavior among households. Strikingly, we also find that larger unconditional grants completely crowded-out the effect of incentives. Our findings suggest that performance incentives can be effective in bureaucratic environments – but also that discretionary resources can fully crowd-out their effect.

Renfu Luo
Chinese Academy of Sciences
luorf.ccap@igsnrr.ac.cn

Grant Miller
CHP/PCOR
Stanford University
117 Encina Commons
Stanford, CA 94305-6019
and NBER
ngmiller@stanford.edu

Scott Rozelle
Stanford University
Freeman Spogli Institute for International Studies
Encina Hall E407
Stanford, CA 94305-6055
rozelle@stanford.edu

Sean Sylvia
School of Economics
Renmin University of China
No 59 Zhongguancun Ave
Beijing, China 100872
sean.sylvia@gmail.com

Marcos Vera-Hernández
University College London
Department of Economics
Gower St.
London WC1E 6BT
United Kingdom
m.vera@ucl.ac.uk

A randomized controlled trials registry entry is available at
http://www.povertyactionlab.org/Hypothesis-Registry
An online appendix is available at
http://www.nber.org/data-appendix/w21302

1. Introduction

The provision of public services in many developing countries is poor (World Bank 2004; Banerjee, Deaton, and Duflo 2004; Das, Hammer, and Leonard 2008; Berendes et al. 2011). 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. Absenteeism in many parts of the world is pervasive (Chaudhury and Hammer 2004; Kremer et al. 2005; Chaudhury et al. 2006; Lewis 2006; Banerjee and Duflo 2006), and providers often fail to do in practice what is within their knowledge and means (Das and Hammer 2004; Alcázar et al. 2006; Chaudhury et al. 2006; Das and Hammer 2007; Leonard and Masatu 2010; Das et al. 2012; Sylvia et al. 2014).

To better align provider incentives with social objectives, performance pay has become increasingly common in public sector service delivery (Oxman and Fretheim 2008; Eichler and Levine 2009; Miller and Babiarz 2013). Drawing on the logic of performance pay in human resource management (Lazear 1995; Hall and Liebman 1998; Lazear 2000), this approach provides direct financial rewards for achieving pre-specified performance targets. Despite their growing prominence, however, many important conceptual questions about the appropriate use of performance pay in public service delivery remain unexplored (Miller and Babiarz 2013).

A fundamental issue is that although performance pay in the private sector focuses on managers (corporate executives, for example), remarkably little is known about performance pay for managers in bureaucracies.[2] Previous studies of public sector

[2] On performance pay for corporate executives and private sector managers in developed countries, see Jensen and Murphy (1990), Hall and Liebman (1998), Murphy (1999), Hall and Murphy (2003), and Oyer

performance pay largely examine incentives to increase effort among front-line service providers, including teachers and health workers.[3] Performance incentives for managers have broader aims – including greater productive and allocative efficiency with the resources under their control (Holmstrom and Ricart i Costa 1986; Athey and Roberts 2001; Bandiera, Barankay, and Rasul 2007). In bureaucratic environments in which managers are strongly motivated by career concerns (and are rewarded little for good results – but may suffer substantial career harm for bad ones), performance incentives may be vital to encouraging innovation in service delivery (Holmstrom 1982; Holmstrom and Ricart i Costa 1986; Dasgupta and Sarafidis 2009).[4]

Given the scope of their responsibility, a key feature of performance pay for managers is that it may be appropriate to base rewards on outputs of production (linking the compensation of corporate executives to their firm's share value, for example) (Holmstrom 1982; Athey and Roberts 2001).[5] In contrast to performance pay rewarding inputs into production, rewarding outputs strengthens incentives for managers to draw on contextual knowledge about the optimal allocation of organizational resources (improving allocative efficiency), to use existing resources/inputs more efficiently (improving productive efficiency) – and more generally, to innovate. Moreover, for

---

and Shaefer (2005). Bandiera et al. (2007) and Bloom et al. (2011) study these virtues of performance pay for private sector managers in developing countries.

[3] Several recent studies have examined performance pay provided as personal income to front line workers in the health and education sectors, including Lavy (2002), Lavy (2009), Glewwe et al. (2010), Muralidharan & Sundararaman (2011), Duflo et al. (2012), Ashraf et al. (2014) and Behrman et al. (2015). Other recent studies have focused on incentives to institutions paid as budget revenue. These include Bloom et al. (2006), Basinga et al. (2011), Gertler and Vermeersch (2012), Olken, Onishi, and Wong (2014), and Yip et al. (2014). An exception is Rasul and Rogger (2013), who study incentives for bureaucrats in the Nigerian civil service. Behrman et al. (2015) also study incentives for school administrators, but bundled with incentives to students and teachers.

[4] An additional advantage of rewarding managers (rather than front-line workers) is that institution-level outcomes may be cheaper to measure than those attributable to individual workers.

[5] Front line workers have less control over productive outputs – and may therefore be more strongly incentivized by performance pay rewarding the use of pre-specified productive inputs.

outputs jointly produced with service beneficiaries or recipients (health and educational outcomes also depend on household inputs), output-based incentives may motivate agents to prevent offsetting compensatory behavior among beneficiaries (Leonard 2003). In the health sector, almost no previous research of which we are aware examines the effects of performance incentives for actual outputs of health production: good health.[6]

Although interest in public sector performance pay aims to improve the efficiency of public expenditures, there is also little evidence about how performance incentives interact with the size of operating budgets.[7] Bureaucracies in developing countries generally operate under stringent budget constraints – and increasing the size of their budgets is often considered an equivalent priority. Importantly, incentives and resources available to managers may function as either complements or substitutes in public service delivery. Complementarity is possible for intuitive reasons – for example, managers may wish to respond to performance incentives, but resources may facilitate their ability to do so.[8] There are also theoretical rationales for substitution – if the marginal return to effort devoted to improving efficiency (reducing waste) declines when managers have more resources at their disposal, managers may simply substitute resources for their own effort.[9]

---

[6] The two exceptions of which we are aware are Singh (2011) and (Miller et al. 2012). Leonard (2003) studies traditional healers' use of outcome-contingent contracts in Cameroun. In the education sector, performance pay rewarding good test scores is more common (Lavy 2009; Muralidharan and Sundararaman 2011; Behrman et al. 2015).

[7] Besley and Ghatak (2007) note that interest in applying performance pay to the public sector (and the philosophy underlying "New Public Management" more generally) was "born out of efforts to decrease the size of public funds going to public goods and services while preserving service levels" in the UK.

[8] Some research has emphasized the connection between incentives and control in organizations (for example, see Prendergast (2002)).

[9] Studying public procurement in Italy, Bandiera et al. (2009) estimate that 83% of total waste in public service provision is attributable to "passive waste," which occurs (in part) because of weak incentives to minimize cost.

More generally, performance incentives may be ill-suited for public sector managers altogether – because other institutional incentives may dominate them or because well-known concerns with performance pay may apply with greater force, for example. First, career concerns can be particularly strong in bureaucracies – and may overpower incentives created by performance pay (Gibbons and Murphy 1992). Second, common concerns about multi-tasking may be exacerbated by the fact that bureaucracies often have a large number of ill-defined tasks and responsibilities (Dewatripont, Jewitt, and Tirole 1999; Dixit 2002; Besley and Ghatak 2007). Third, compared with their private sector counterparts, civil servants may be more intrinsically or pro-socially motivated (Francois 2000; Francois and Vlassopoulos 2008; Tonin and Vlassopoulos 2014) – and performance pay may dampen the effects of intrinsic motivation (Deci and Ryan 1985; Fehr and Falk 2002; Bénabou and Tirole 2006; Francois and Vlassopoulos 2008; Gneezy, Meier, and Rey-Biel 2011; Kamenica 2012).

In this paper we study these issues directly, analyzing a large-scale policy experiment in rural China rewarding primary school administrators (head principals, hereafter "administrators") – managers and executive decision-makers in their schools – for improvements in the health of their students. Specifically, we offer several sizes of financial incentives to administrators for measurable reductions in student anemia – a leading child health problem in rural China – over the course of an academic year.[10] We

[10] Previous studies have shown anemia rates among primary school aged students in poor regions of western China to be around 30% on average (Luo et al. 2011; Miller et al. 2012).

also analyze how the response of managers to performance incentives varies with randomly assigned block grants.[11]

Our study yields four key findings. First, we find that larger incentives for anemia reduction were effective when administrators had fewer resources at their disposal for implementing the program. Incentives that provided substantial additional income to administrators (mean realized payouts of about 2 months of annual salary) reduced anemia among students anemic at baseline by 13.8 percentage points (or 38%). Importantly, we also find that directly rewarding outputs (good health) in this way led school administrators to innovate in working with students' parents to improve diets at home (as well as at school). This finding suggests that outcome-based performance incentives may be an effective tool to mitigate potentially offsetting compensatory behavior among beneficiary households in response to supply-side investments in child education and health.[12]

Second, we find that small incentives (one tenth the size of the larger incentives) were ineffective in reducing anemia – and were significantly less effective than large incentives. A number of studies outside of organizational settings have shown that even small incentives (or the existence of incentives alone, independent of size) can lead to meaningful changes in behavior. Several studies report large increases in incentivized behavior in response to modest rewards (as well as highly elastic demand at prices close to zero) (Kremer and Miguel 2007; Thornton 2008; Banerjee et al. 2010; Cohen and

[11] Because all administrators in our study are provided with information about anemia and a small block grant, the effect of incentives that we estimate is net of effects due to "labeling" or increasing the salience of anemia (Benhassine et al. 2013).

[12] That school-based inputs can potentially crowd-out household investments in nutrition and education has been noted in a number of contexts (Jacoby 2002; Kazianga et al. 2009; Islam and Hoddinott 2009; Buttenheim, Alderman, and Friedman 2011; Alderman and Bundy 2012; Das et al. 2013). The welfare consequences of mitigating household responses to supply-side investments are context dependent.

Dupas 2010; Karlan, List, and Shafir 2011; Duflo, Kremer, and Robinson 2011). Alternatively, other research suggests that small incentives may *reduce* effort if the potential for financial rewards are dominated by negative effects due to the information conveyed by incentives – for example, if incentives cast doubt that a task is performed for prosocial reasons (Gneezy and Rustichini 2000; Bénabou and Tirole 2006; Gneezy, Meier, and Rey-Biel 2011). Our finding that small incentives do not lead to significant effects in either direction are more consistent with a linear monotonic relationship in the context that we study.[13]

Third, we find that even absent explicit experimental incentives, unrestricted budget transfers to school administrators led to sizeable reductions in anemia. This result suggests that administrators appear motivated to allocate resources to the nutrition program, even absent explicit incentives to do so. However, the resource cost of reducing anemia through larger school budgets was approximately twice as great (per case of anemia averted) as combining larger performance incentives with smaller budgets – implying that school administrators with incentives used smaller budgets with greater productive efficiency.

Finally, whether or not unrestricted grants and incentives are substitutes depends on incentive size. We find that small incentives do not interact with grants, but large incentives and grants are substitutes. The latter effect is particularly strong: unrestricted resource transfers (of sizes chosen by government planners in practice) completely crowd-out the effect of large incentives. As our theoretical model shows, this is possible

[13] We also note at least two key differences with this previous literature: one is that receiving incentive payments in our study required sustained behavior change (reducing iron deficiency anemia requires several months of dietary chance), and another is that because all school administrators receive information about anemia, the presence of incentives may not alter its salience as much (Benhassine et al. 2013).

because administrators may substitute school resources for organizational effort when there are adequate resources to do so. Importantly, we find this pattern of results not only with student nutritional outcomes, but also with intermediate measures of resource allocation and effort (for example, the provision of better nutrition and efforts to persuade parents to improve their children's diets at home). Substitution therefore reflects reductions in administrator effort with larger budgets (and not simply decreasing marginal returns to inputs in the biological production of child nutrition).[14]

As with any empirical study, our results are specific to our study environment, but we make several contributions to the literature. First, the existing handful of studies on performance incentives in public service delivery have focused largely on incentives to front-line workers (see Lavy 2002, 2009; Glewwe et al. 2010; Muralidharan & Sundararaman 2011; Duflo et al. 2012; Ashraf et al. 2014; Behrman et al. 2015).[15] To the best of our knowledge, this study is the first to experimentally analyze the provision of performance incentives to public sector managers. Second, we not only compare the effects of incentives and additional resources (as in Lavy 2002), but we also examine complementarity/substitution between them – as well as how their interaction varies with the size of both incentives and grants. Evidence on these interactions is critical for the design of public sector incentive systems.

The rest of this paper is organized as follows. Section 2 presents a conceptual framework for understanding school administrators' behavioral responses to output-based

[14] In this paper, we study health (specifically, anemia). In a companion paper, Miller et al. (2015), we study test scores and unintended behavioral responses (specifically, multitasking) to incentives for anemia reduction. In short, we find no significant changes in test scores associated with incentives for anemia reduction or with larger block grants.

[15] An exception is Behrman et al (2015) who rewarded school administrators, but in combination with students and teachers rather that on its own. Bloom et al. (2006) and Olken et al. (2014) study incentives paid as operational revenue at the organizational level.

performance incentives. Section 3 provides background on school-based nutrition programs as well as the causes and consequences of anemia. Section 4 describes our experimental design, data collection, and methods. Section 5 reports our results, and Section 6 concludes.

2. Conceptual Framework

In this section, we propose a simple model of the school administrator decision problem that we study. Specifically, we consider the influence of both output-based performance incentives and of discretionary resources on organizational effort – as well as how they interact (if they are substitutes or complements). As in Rogger (2014) and Nath (2015), we model the school administrator (bureaucrat) as taking an effort decision, e, which will be relevant for reducing anemia in the school. In addition, the school administrator also decides on the allocation of resources – in particular, how to divide the school budget, G, between anemia reduction, A, and other school functions, G-A. The level of effort, e, together with the funds allocated to reducing anemia, A, will determine the amount of anemia reducing inputs, i (food and iron supplements) that students effectively consume. Hence, the school administrator will solve:

$$\max_{e,i,A}\ w + \theta H - v(e) + S(G - A) \tag{1}$$

subject to:

$$w = tH + m \tag{2}$$

$$H = h(i) \tag{3}$$

$$i \leq (1 + \alpha e)A \tag{4}$$

$$G - A \ \geq 0 \tag{5}$$

Total take home pay, *w,* consists of base pay, *m*, and a reward for improving student health, *tH* – which is the product of *t*, the marginal bonus, and *H*, the net gain in student health (in our case, the net reduction in the number of students with anemia). *H* is produced with the chosen amount of inputs, *i*, according to the production function *h(i)*, which we assume to be strictly increasing and concave: $h' > 0, h'' \leq 0$. The disutility of effort, *v(e)*, is also strictly increasing but convex: $v' > 0, v'' \geq 0$. The parameter θ, which is non-negative, allows the school administrator to be altruistic, deriving direct utility from student health (pro-sociality and public service motivation are often cited as important among public sector workers – for example, see Besley and Ghatak (2007) and Dal Bó, Finan, and Rossi (2013). The school administrator also derives utility from school functions unrelated to health, *S(G-A)*, which is also assumed to be increasing (*S'>0*) and concave (*S''<0*).

Equation (4) implies that the school administrator can exert effort to use the budget dedicated to reduce anemia, *A*, more efficiently. Specifically, the marginal increase in *effective* resources due to effort is given by $\alpha > 0$. This could reflect effort devoted to making inputs more effective (ensuring that students take multivitamins provided to them, for example), shopping around to acquire inputs at lower prices, or identifying and choosing a strategy that will be most effective given local conditions. Additionally, because student health is jointly produced with households, effort could also focus on increasing the amount of resources by working with parents to increase health and nutrition investments at home. For school functions unrelated to nutrition, *S(G-A),* we abstract away from modeling choice of effort so that our predictions about substitution/complementarity between incentives and resources are not mediated by multitasking.[16]

To solve the model, we use the fact that constraint (4) is binding and hence $e = \left(\frac{i}{\alpha A} - \frac{1}{\alpha}\right)$.[17] The first order conditions with respect to *i* and *A* are given respectively by:

$$(t + \theta)h'(i) = \frac{1}{\alpha G} v' \left(\frac{i}{\alpha A} - \frac{1}{\alpha}\right), \tag{6}$$

and

$$\frac{i}{\alpha A^2}\, v' \left(\frac{i}{\alpha A} - \frac{1}{\alpha}\right) = S'(G - A). \tag{7}$$

The first order condition, (6), implies that the school administrator chooses the amount of health improving input i by equating its marginal benefit (the increase in health, $h'(i)$, multiplied by $t + \theta$, reflecting both the increase in take home pay and the altruistic increase in direct utility) with its marginal cost. The first order condition, (7), implies that resources G are invested in activities unrelated to nutrition up to the point that its marginal product equals the marginal cost of effort to keep i at its optimal level (the fewer resources dedicated to nutrition, the higher effort must be in compensation to maintain i at its optimal level).

For what follows, it is useful to write the objective function (1) in terms of *A* and *i*, making explicit that the optimal effort is a function of *i* and *A*, $\tilde{e} = \left(\frac{i}{\alpha A} - \frac{1}{\alpha}\right)$. Then the first derivatives are:

$$V_i' = (t + \theta)h'(i) - \frac{d\tilde{e}}{di} v'(\tilde{e}), \tag{8}$$

[16] Admittedly, the effort exerted in non-health functions could decrease if the effort in anemia reduction increases, and the cross-partial derivative of the effort cost function is positive.
[17] It is useful to derive predictions in terms of *i* rather than *e* because it is what we observe in the data.

$$V_A' = -\frac{d\tilde{e}}{dA}\, v'(\tilde{e}) - S'(G-A), \tag{9}$$

and the second derivatives are:

$$V_{ii}'' = (t+\theta)h''(i) - \left(\frac{d\tilde{e}}{di}\right)^2 v''(\tilde{e}) < 0, \tag{10}$$

$$V_{AA}'' = -\left(\frac{d\tilde{e}}{dA}\right)^2 v''(\tilde{e}) - \left(\frac{d^2\tilde{e}}{dA^2}\right) v'(\tilde{e})\, S''(G-A) < 0, \tag{11}$$

$$V_{iA}'' = \left(\frac{d\tilde{e}}{dA}\right)\left(\frac{d\tilde{e}}{di}\right) v''(\tilde{e}) - \left(\frac{d^2\tilde{e}}{didA}\right) v'(\tilde{e}) > 0, \tag{12}$$

$$V_{AG}'' = -S''(G-A) > 0. \tag{13}$$

The second order condition for a maximum require that: (1) $V_{ii}'' < 0$, and $V_{AA}'' < 0$, which is trivially satisfied given the concavity assumptions made, and that

(2) $|SOC_2| = (V_{ii}'')(V_{AA}'') - (V_{iA}'')^2 > 0.$ (14)

2.1 Comparative Statics

Our main interest is to analyze how the school administrator's use of health inputs changes with discretionary resources, *G*, and incentives, *t* – both separately and in combination (as we study in our experiment). The administrator's health input use changes with discretionary resources (*G*) according to the expression:[18]

$$\frac{di}{dG} = \frac{\begin{vmatrix} -V_{ig}'' & V_{iA}'' \\ -V_{AG}'' & V_{AA}'' \end{vmatrix}}{|SOC_2|} = \frac{\begin{vmatrix} 0 & V_{iA}'' \\ -V_{AG}'' & V_{AA}'' \end{vmatrix}}{|SOC_2|} = \frac{V_{AG}''V_{iA}''}{|SOC_2|} > 0, \tag{15}$$

which is positive: the more resources the school has the more anemia reduction inputs will be acquired. The change in input use with incentives (*t*) is given by:

[18] For simplicity, in what follows we omit the functions' arguments, which are the same as in equations (6) to (13).

$$\frac{di}{dt} = \frac{-V''_{it}V''_{AA}}{|SOC_2|} = \frac{h'(i)\left[\left(\frac{d\tilde{e}}{dA}\right)^2 v''(\tilde{e}) + \left(\frac{d^2\tilde{e}}{dA^2}\right)v'(\tilde{e}) - S''(G-A)\right]}{|SOC_2|} > 0, \tag{16}$$

which is also positive.

Equation (16) is also useful for understanding how the administrator's response to incentives will vary with resources. Focusing on the numerator of (16), if G is relatively large, so A is, together with i, then h', v', and v'' will be small – and hence the effect on the incentive on the inputs will also be smaller. More formally, the cross partial derivative between incentives and discretionary resources is given by:

$$\frac{d^2 i}{dGdt} = \frac{1}{|SOC_2|^2}\left[S''(G-A)V''_{iA}\, h''(i)\, V''_{AA}\right] < 0, \tag{17}$$

which is negative – implying that discretionary resources and incentives are substitutes. The intuition of why the marginal effect of incentives decreases with additional resources is as follows. Additional resources are split between nutrition and other unrelated tasks. Of the resources devoted to nutrition, part is used to acquire more anemia-reducing inputs, and the remainder is used to reduce the level of effort required to improve the efficiency of nutrition spending. Hence, the marginal effect of the incentives is smaller.

Note, however, that discretionary resources and incentives might be complements in models with non-convexities – for example, if greater resources increase the administrator's return to effort in student health production, for instance, if there is a technology with fixed cost $F > 0$ that enhances the productivity of effort. An example of such technology is a refrigerator that allows one to store meat bought in the market, reducing time and money spent on trips to the market. The administrator could choose to buy this technology, which would imply that Equation (4) becomes $i \leq (1 + \beta e)(A -$

$F)$, with $\beta > \alpha$. Incentives and block grants would be complements if G increased sufficiently, and β is sufficiently larger than α.

3. Background

3.1 School-Based Nutrition 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 and Guyatt 1996; Jukes, Drake, and Bundy 2008; Orazem, Glewwe, and Patrinos 2008). Because schools are natural points of contact with school-aged children, they may provide a platform from which health and nutrition interventions can be delivered at relatively low cost (Bundy and Guyatt 1996; Bundy et al. 2006; Jukes, Drake, and Bundy 2008). Because of this, school-based health, nutrition and feeding programs are a ubiquitously[19] central function of schools, particularly in developing countries. In China, schools have the legal responsibility to promote the health of their students (Education Law of the Peoples Republic of China, 1995).

3.2 The Causes and Consequences of Anemia

Our study examines school-based programs to reduce anemia. Anemia is estimated to affect nearly one quarter of all school-aged children worldwide (World Health Organization 2001). Although there are many causes of anemia (including a variety of genetic disorders and infections as well as nutritional deficiencies), iron

[19] Nearly every country in the world feeds students through school-based programs (Alderman and Bundy 2012).

deficiency accounts for about 50% of cases globally (Balarajan et al. 2011; Pasricha et al. 2013)[20] – and 85-95% of cases in China (Du et al. 2000).

The consequences of iron deficiency—with or without anemia—can be substantial, particularly for children at critical stages of development. A large literature links iron deficiency to fatigue and reduced work capacity among adolescents and adults, impaired cognition and cognitive development among children, and reduced immune response for all age groups (Thomas et al. 2006; R. Yip 2001; World Health Organization 2001; Balarajan et al. 2011).[21] School-aged children with anemia (the focus of our study) have also been shown to have inferior educational outcomes (grades, attendance, and school attainment – Taras 2005; Nokes, van den Bosch, and Bundy 1998).

3.3 Biomedical Strategies for Reducing Anemia

Increasing iron consumption can effectively prevent iron deficiency anemia. Worldwide, fortifying staple foods with iron has historically been an effective approach to addressing micronutrient deficiencies (Allen et al. 2006). Fortification is an attractive strategy because it requires little behavior change and because it can be implemented on a large scale. However, fortification of staple foods may be ineffective in areas like Northwest China in which households grow and consume their own food (Allen et al. 2006).

An alternative approach is to increase the consumption of naturally iron-rich foods and those that promote iron absorption during digestion. Animal sources (including

---

[20] 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).

[21] 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).

red meats, fish, and poultry) provide *heme* iron, which is more easily absorbed during digestion; plant sources (including green, leafy vegetables) provide *non-heme* iron, which is less readily absorbed – but can be promoted by consumption of vitamin C (and inhibited by consumption of milk and other calcium-rich products).[22]

Finally, a third approach is the delivery of micronutrient supplements (for example, vitamins) containing iron. To be effective, however, regular consumption over several few months is necessary – and so inadequate compliance may render supplementation ineffective (Bobonis et al. 2006; Bhutta et al. 2013; Pasricha et al. 2013; Martorell et al. 2015).[23]

4. The Experiment

4.1 Sampling

To draw our study sample, we began with all 36 counties officially designated by the Chinese government as "poverty counties" in five regions (prefectures) in western China (Haidong in Qinghai Province, Dingxi, Tianshui, and Longnan in Gansu Province, and Ankang in Shaanxi Province – see Figure 1). In August 2011, we conducted a canvass survey in each county to construct a list of all rural primary schools and the number of students enrolled in each. Restricting our sampling frame to primary schools

[22] 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.

[23] 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.

with 150-300 students total,[24] we randomly selected 170[25] of 1,410 eligible schools for inclusion in our study (and limited our selection to one school per township[26]). Our sample size was based on power calculations conducted using data from primary schools in the same region of China (Miller et al. 2012).[27]

Within study schools, we randomly sampled 50 fourth and fifth grade students from each school. In China, fourth and fifth grade students are typically 10 to 11 years old, and we chose these grades to select students whom we considered sufficiently old to provide meaningful survey responses – but also sufficiently young to be generally pre-pubescent (given the independent effect of menarche on hemoglobin concentration).[28] We also conducted physical exams and collected data from students from other grades at baseline to obfuscate our focus on fourth and fifth graders.

4.2 Data Collection

We conducted our baseline survey in September 2011 and our follow-up survey in May 2012 (at the beginning and end of the 2011-2012 academic year), collecting detailed information on students, households, school administrators, and schools.

---

[24] 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.

[25] The full experiment involved an additional 130 schools, which were allocated to two other experimental groups where school administrators were offered incentives to based on test scores and "dual" incentives based on both anemia reduction and test score improvement. The results for these additional study arms are reported in a companion paper (Miller et al. 2015).

[26] 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.

[27] Using data from Miller et al. (2012), we performed Montecarlo simulations to conduct power calculations for students who were anemic at the time of that study's baseline survey. The intra-class correlation was estimated adjusting for covariates (baseline hemoglobin concentration, the number of students in each school, whether schools had a kitchen, student-teacher ratio, distance to the furthest village served by the school, percent boarding students, and county dummies), which we also specified as covariates in the current study's pre-analysis plan.

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

Student Surveys. We interviewed all sampled students at their school, collecting information on student background, health behaviors related to anemia, school activities, and general health. To collect information on school and home feeding practices, students were also given standard food frequency questionnaires to record information about food consumption at school and at home over the past week.[29]

We also measured student blood hemoglobin (Hb) concentration at the time of the student survey. Nurses from the Medical School of Xi'an Jiaotong University accompanied study enumerators, collecting finger-prick blood samples to analyze on-site (at schools) using HemoCue Hb 201+ assessment systems.

Household Surveys. For each sampled student, we also collected information on students' households using forms completed by parents.[30] Specifically, these surveys collected information about interactions between parents and the school, household income and assets, health-related expenditures, expenditures on food and information on other household members, focusing on household characteristics that students would be unlikely to know themselves.

School Administrator Surveys. We interviewed school administrators (bureaucrats) at three different points in time: before and after school administrators were told about the incentive contract and block grant to which they were assigned and again at endline. At baseline, school administrators provided information about their

[29] Information on food consumption was collected using a seven-day recall "food frequency questionnaires" (FFQs) completed 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. Food frequency questionnaires (FFQs) 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). FFQ responses by children about their own consumption has been shown to be more accurate than the responses of their parents (Burrows et al. 2013).

[30] For budgetary reasons, household surveys were given to students to take home and return. As a result, household forms are missing for approximately 20% of students. All possible information on students and households was collected with the student survey, which was administered by enumerators.

background, job history, salary, and compensation as well as perceptions of professional responsibilities and anemia knowledge. Using scales adapted from Grant (2008), we also measured the intrinsic and pro-social motivation of administrators. Following their participation in the training session on anemia (conducted 3 weeks after the baseline survey) administrators were given a second short survey to measure their understanding of the training material.

School Surveys. Finally, we collected basic information from schools (about enrollment, staffing, facilities, finances, and meal provision) and teachers (about teacher characteristics, communication with parents, and teaching practices).

4.3 Experimental Design

We designed our study as a cluster-randomized trial using a 3×2 crosscutting design (Figure 2). After conducting our baseline survey, we provided all school administrators with information about anemia (see our written materials in the online appendix – and which also included a video presentation by a Chinese nutrition specialist), and schools were randomly assigned one of six experimental cells (see Figure 3 for the study timeline). The first three paths of Figure 2 show randomly-assigned incentive groups: a group without incentives (Group A), a "small" incentive group (Group B), and a "large" incentive group (Group C). Across these arm are two orthogonally-assigned block grant groups: a "small" block grant group (Group 1) and a "large" block grant group (Group 2). The reference group in our six-cell design is the default policy (education about anemia coupled with a modest resource transfer and no incentives, Group A1).[31]

[31] In a previous study, we find that educating school administrators on anemia (including the same information as in the current study) alone, without incentives or grants, had no detectable impact on anemia

To improve power, we used a stratified randomization procedure. Specifically, using joint quintiles of the baseline distribution of school-level hemoglobin concentration and combined standardized math and Chinese exam scores – yielding 25 strata, we randomized cell assignment within each stratum. Our analysis takes this randomization procedure into account, conditioning on stratum fixed effects (Bruhn and McKenzie 2009)

Incentives for Anemia Reduction. In the large incentive arm (65 schools, Group C in Figure 2), we offered school administrators financial incentives to be paid as private income according to the net reduction in number of students identified as anemic between the beginning and end of the school year. The specific structure of the large incentive contract was:

$$P = \begin{cases} 125\ RMB * (N_b - N_e) & if\ (N_b - N_e) > 0 \\ 0 & otherwise \end{cases}$$

where N_b is the number of students found to be anemic at baseline and N_e is the number of who were anemic at the time of the endline survey.[32] Based on an earlier study (Miller et al. 2012), the contract increment (125 *yuan* (RMB), or about \$19.40[33]) per student reduction was chosen to provide roughly two months of a school administrator's annual salary for a feasible reduction in anemia given previous studies (a reduction of about 50%).[34] Actual payouts for school administrators with the large incentive and small

rates (Miller et al. 2012). Our reference group also mimics how a recent Chinese school nutrition program (costing 16 billion yuan per year) was designed. 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.

[32] We measured anemia using a sample of 50 4th and 5th graders and calculated the implied number of anemic children in the school using the prevalence rate in our sample.

[33] We use a conversion rate of \$1 = 6.4 RMB, the approximate exchange rate at the time of the baseline survey (September 2011).

[34] There are presumably superior contract structures, but optimal contract design requires substantial information not available to us, including information about the cost of provider effort, the productivity of

block were ultimately 3,303 *yuan* (or about $516) – approximately two month's base pay for school administrators in this region. We did not reveal the identity of students who were anemic at baseline to administrators (and when we asked teachers to identify students who were anemic at endline, they were unable to do so).[35]

The small incentive arm (40 schools, Group B in Figure 2) was identical to the large incentive arm except that the magnitude of the incremental incentive was ten times smaller (12.5 RMB, or about $1.95 per student reduction in anemia between baseline and follow-up in our sample). This magnitude of this incentive provides roughly 0.2 additional months of annual salary for the same feasible reduction in anemia given previous studies.

At the time that school administrators signed incentive contracts, they were told the (implied) number of anemic students in their schools (the identity of anemic children was not revealed).[36] Contracts were written using official letterhead of the Chinese Academy of Sciences (a government agency) and counter-signed by the deputy director of the implementing research center (school administrators signed two copies of the contract, one of which they kept). Note that all interventions were implemented in partnership with local education bureaus, signifying to school administrators that the project was sanctioned by local governments.

Block Grants. The small block grant (Group 1 in Figure 2) was 0.3 RMB ($0.05) per student per day (85 schools), which we calculated to be adequate for school

provider effort, and the utility functions of both providers and the contracting 'principal' (Laffont and Tirole 1993; Salanié 2005). Simple, easily understandable contracts may also appear more transparent to school administrators and promote credibility.

[35] We did reveal the identity of students who were severely anemic (with hemoglobin concentration below 80 g/L) as these students required immediate medical attention. There were 3 such students found at baseline.

[36] Note that administrators in all study cells were provided the same information about the number of anemic children in their respective schools.

administrators to purchase vitamins for each student to take daily. The large block grant (Group 2 in Figure 2) was 0.7 RMB ($0.11) per student per day (85 schools). In total, small block grant schools received 7,452 yuan ($1,164) on average and large block grant schools received 17,388 *yuan* ($2,717). These grants were given to schools in two installments, once at the beginning of the program and another approximately half way through the school year.[37] Although funds were given in the context of the nutrition program roll-out, administrators were explicitly told that they were free to allocate these to other school functions at their discretion.[38] Indeed, Figure 4 shows that administrators used a substantial share of their grants for activities unrelated to nutrition.

Health Education. Because knowledge about anemia in our study areas was poor, prior to revealing treatment assignment, we provided health education about nutrition and anemia to all school administrators in our study (see online appendix for materials). Our health education materials were based on published, peer-reviewed studies and specifically included information about: 1) the prevalence and causes of anemia, 2) the consequences of anemia (including its effect on cognitive development and academic performance), and 3) 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.).

4.4 Balance and Attrition

[37] After explaining block grant assignment to administrators, we asked them to complete a non-binding budget plan for how they intended to use the block grant. Our study team emphasized that this plan was non-binding, but this plan would be used to coordinate orders for iron supplements to be delivered to schools. This was necessary because the market for supplements in rural areas is limited. Administrators were free to change their supplement orders at any time.

[38] Note that while these transfers were not large compared to total school expenditures, they do represent a significant increase in budgetary autonomy for school administrators as the bulk of school expenditures are earmarked for specified uses at higher levels of administration.

Summary statistics and tests for balance across study arms are shown in Table 1.[39] Panel A shows student level characteristics (N=2051), Panel B shows characteristics of schools (N=167), and Panel C shows characteristics of school administrators (N=167).[40] The first two columns of the table give the mean and standard deviation of each variable in the comparison (small block grant, no incentive) group. Columns (3) – (7) show coefficients on treatment variables and interactions estimated using Equation (18), controlling only for randomization strata fixed effects. The final column shows the p-value from a test that the coefficients are jointly zero for each characteristic. Only 4 of the 75 tests are significant at the 10% level, and a test for joint equality is rejected at the 10% level for only one characteristic (the number of times meat was consumed in the past week).[41] Joint tests for all 15 characteristics reveal no significant differences.[42]

The overall attrition rate between baseline and endline surveys was 6.2% in our sample of children anemic at baseline (5% for the full sample). Defining attrition as a missing hemoglobin measurement at endline for students with a baseline measurement, Appendix Table 2 shows that there were no meaningful differences in attrition across treatment groups (Columns 1 & 2). Analyzing the correlates of a missing household survey at endline conditional on a child not dropping out, Appendix Table 2 also shows

---

[39] This table shows summary statistics and tests balance for our main analysis sample of students initially anemic at baseline. Summary statistics and balance tests for the full sample are given in Appendix Table 1.

[40] Although 170 schools were included in the study, no students were found to be anemic in 3 schools at baseline. No schools refused participation in the study. The baseline anemia rate (defined as Hb<120 g/L) in the full sample was 24%.

[41] Because baseline differences in meat consumption could possibly affect results, we repeated all analyses controlling for baseline meat consumption. There were no significant differences in results and very minor differences in point estimates (for example, the effects we estimate for the large anemia incentive and large grant on anemia change less than 0.1 percentage points). These results are available upon request.

[42] These tests were conducted by regressing treatment status on all 15 baseline covariates and testing that the coefficients were jointly zero. The smallest p-value from these F-tests was 0.29.

that neither the treatment indicators nor other covariates are significantly correlated with a missing household survey form.

4.5 Empirical Strategy

Given random assignment of schools to treatment cells shown in Figure 2, comparisons of outcome variable means across treatment groups provides unbiased estimates of the effect of each experimental treatment. However, to increase power (and to account for our stratified randomization procedure), we condition our estimates on a set of covariates used in power calculations. With few exceptions, all of the analyses presented (including outcome variables, regression specifications, and hypotheses tested) were pre-specified in a pre-analysis plan written and filed before endline data were available for analysis.[43] In reporting results below, we explicitly note analyses that deviate from the pre-analysis plan.

As specified in advance, we use ordinary least-squares (OLS) regression to estimate the effect of cell assignment on child-level outcomes with the following specification:

$$Y_{i,j} = \alpha + \beta_1 SI_j + \beta_2 LI_j + \beta_3 LG_j + \beta_4 (SI_j) \times (LG_j) + \beta_5 (LI_j) \times (LG_j) + X'_{i,j}\gamma + \varepsilon_{i,j} \quad (18)$$

where $Y_{i,j}$ is the outcome for child i in school j; SI_j is a dummy that equals 1 if the administrator in school j was assigned to receive a small incentive contract and 0 otherwise; LI_j is equal to 1 if the administrator in school j was assigned to receive a large anemia reduction incentive contract; LG_j is equal to 1 if the school received a large block grant; $X_{i,j}$ is a vector of child controls (age, class-year, and gender, and baseline value of the outcome variable), school controls (number of students, student-teacher ratio, whether

[43] This analysis plan was filed with the Abdul Latif Jameel Poverty Action Lab at http://www.povertyactionlab.org/Hypothesis-Registry.

the school has a kitchen, proportion of boarding students, and distance to the farthest village in the school's catchment area); and dummy variables for counties and randomization strata. We adjusted our standard errors for clustering at the school level using the cluster-corrected Huber-White estimator.

In addition to estimating effects on our two primary outcomes (hemoglobin concentration and a dichotomous indicator for anemia status), we use the same specification to estimate effects on secondary outcomes to examine the behavioral mechanisms underlying changes in primary outcomes. For these secondary outcomes, we focus our analysis on summary indices constructed using groups of closely-related outcome variables (as we specified in advance). To construct these indices, we used the GLS weighting procedure described by Anderson (2008). For each individual, we constructed a variable $\bar{s}_{ij}$ as the weighted average of k normalized outcome variables in group (y_{ijk}). The weight placed on each outcome variable is the sum of its row entries in the inverted covariance matrix for group j such that:

$$\bar{s}_{ij} = \left(\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_j^{-1}\mathbf{1}\right)^{-1}\left(\mathbf{1}'\widehat{\boldsymbol{\Sigma}}_j^{-1}\boldsymbol{y}_{ij}\right)$$

where $\mathbf{1}$ is a column vector of 1s, $\widehat{\boldsymbol{\Sigma}}_j^{-1}$ is the inverted covariance matrix, and $\boldsymbol{y}_{ij}$ is a column vector of all outcomes for individual i in group j. Because each outcome is normalized (by subtracting the mean and dividing by the standard deviation in the sample), the summary index, $\bar{s}_{ij}$, is in standard deviation units. In addition to reducing the number of tests required, this weighting procedure can improve efficiency by placing less weight on outcomes that are highly correlated and more weight on those less correlated. The summary index variable can also be created for individuals with a subset of missing outcomes (these outcomes simply receive less weight in the construction of the index).

Although we emphasize these indices in our discussion, we also report estimates for each individual index component in Appendix Tables 5 to 8.

A note on correcting for multiple comparisons is also warranted. For our primary outcomes, we test eight null hypotheses: five for treatment main effects and their interactions (shown in Equation 18)) and three additional ones – that the small and large incentives have the same average effect ($\beta_1 = \beta_2$), that the large incentive and the large block grant have the same average effect ($\beta_2 = \beta_3$), and that the average effect of the large incentive in presence of a large grant is zero ($\beta_2 + \beta_5 = 0$).[44] We therefore adjust our p-values to control the Family Wise Error Rate (FWER). Specifically, we use the free step-down resampling method of Westfall and Young (1993), which accounts for the dependency of the data. For secondary outcomes, we adjust our p-values according to the total number of tests within a family of outcomes (the number of outcomes in the family times five – the number of treatment coefficients in each regression).

5. Results: Childhood Anemia and Underlying Behavioral Responses

In this section, we first present results obtained by estimating Equation (18) for anemia status and hemoglobin concentration, and in Section 5.2, we then investigate the underlying behavioral responses that may have produced them. Following our pre-analysis plan, we emphasize estimates from our sub-sample of children who were anemic at baseline. In the Appendix we report results for the full sample of children receiving hemoglobin tests (Appendix Tables 3, 4, 6, and 8).

5.1. Childhood Anemia

[44] We did not pre-specify the last of these (whether or not the joint effect of the large grant and large incentive is negative), but we did pre-specify whether or not incentives and large grants are complements or substitutes ($\beta_4 = 0,\ \beta_5 = 0$).

The first five rows of Table 2 report estimates for each treatment and their interactions (and the seventh row reports comparison group means for the no incentive, small grant group at endline). For each estimate, we report the regression coefficient, the standard error and corresponding p-value, and the p-value adjusted for multiple hypotheses testing.

Result 1: Large Incentives. First, we find that the large incentive significantly reduced the probability of anemia at endline. Specifically, the large incentive was associated with a 14 percentage point reduction in anemia (Table 2, Row 2, Column 1; unadjusted p-value=0.001, adjusted p-value=0.064), implying a 37.9% reduction relative to the comparison group (small grant, no incentive schools) at endline. The corresponding increase in hemoglobin was about 2.6 g/L (Table 2, Row 2, Column 2; unadjusted p-value=0.015, adjusted P-value=0.285). These empirical findings agree with our model's prediction derived in Equation (16).

Because our incentives rewarded anemia reduction (and not hemoglobin levels per se) and anemia status reflects shifts in the distribution of altitude adjusted hemoglobin concentrations across the 120g/L threshold, Figure 5A plots the distribution of endline hemoglobin concentrations (adjusted for covariates included in Equation 18) by study arm among children who were anemic at baseline. The distribution for the large incentive group is shifted to the right of the control group distribution (Kolmogorov-Smirov test p-value=0.02). This relative shift in mass is greater in the left tail of the distribution, implying that the large incentive reduced the share of children falling below the anemia threshold.

Result 2: Small Incentives. Second, in contrast, the small incentive had no detectable effect on the probability of anemia at endline (Table 2, Row 1, Column 1).[45] Comparing the estimates for small and large incentives ($\beta_1 = \beta_2$ in Equation (18)), we also reject the null hypothesis that the two estimates are equal (Table 2, Row 8, Column 1; adjusted p-value=0.089). Taken together, the estimates for the small and large incentives suggest that the price effect of incentives is meaningful independent of information conveyed by the presence of an incentive contract (Gneezy and Rustichini 2000). Figure 5A shows that the shift in the hemoglobin distribution for the small incentive arm relative to the control group arm is smaller – particularly in the left tail of the distribution.

An important question in the literature on financial incentives is whether or not they crowd-out intrinsic or pro-social motivation (Deci and Ryan 1985; Gneezy and Rustichini 2000; Fehr and Falk 2002; Francois and Vlassopoulos 2008; Gneezy, Meier, and Rey-Biel 2011; Kamenica 2012). We find that the small incentive increased the anemia rate by 31 percentage points (adjusted p-value= 0.038) amongst school administrators who score higher at baseline on a pro-sociality scale (adapted from Grant (2008)). We also find a similar effect for intrinsic motivation (also adapted from Grant (2008)), but it is not statistically significant (coefficient = 0.15, adjusted p-value = 0.60). However, the effect of the large incentive is not heterogeneous by pro-social or intrinsic motivation (the coefficients are close to zero and not statistically significant), implying that if monetary incentives are large, crowding-out of pro-social motivation may be overcome by extrinsic motivation provided by incentives.

[45] Note that although the number of schools in the small incentive group is smaller, the standard errors in the small and large incentive arms are similar, but the small incentive coefficient is close to zero (-0.012).

Result 3: Large Block Grants. Third, in the absence of any explicit incentive, the large block grant alone reduced the probability of student anemia at endline (an unambiguous prediction of our model, as Equation 15 shows). Specifically, Table 2 (Column 1, Row 3) shows that the large block grant was associated with a 14.5 percentage point reduction in anemia (adjusted p-value=0.047), implying a 39.8% reduction relative to the comparison group at endline. This reduction is very similar to the effect of the large incentive (-0.145 vs. -0.138), but the average increase in hemoglobin concentration is larger (4.205 vs. 2.567), although not statistically so (Table 2, Row 9, Column 2; adjusted p-value=0.597).

Result 4: Interactions between Incentives and Grants Depend on Incentive Size. Whether or not incentives and unrestricted grants are complements or substitutes is an empirical issue. The model in Section 2 predicts that incentives and block grants will be substitutes (unless there are investments requiring substantial resources, but that substantially enhance the productivity of effort – which would produce complementarity). We do not find evidence of complementarity – and notably, incentives and block grants can be strong substitutes if the incentives are sufficiently large.

Specifically, when incentives are small, they do not interact with the large grant. Table 2 (Column 1, Row 4) shows that the estimate for the interaction between the small incentive and the large grant (β_4 in Equation 18) is not statistically different from zero (p-value=0.888) in their impact on anemia status. However, the interaction between the large incentive and the large block grant (β_5 in Equation 18; Table 2, Column 1, Row 5) is positive and statistically significant (adjusted p-value=0.072). Moreover, the magnitude of substitution implies that the large incentive and the large block grant fully crowd each other out: the marginal effect of the large incentive given the large block grant in Column 1 ($\beta_2 + \beta_5 = 0.058$) is not statistically different from zero (adjusted p-value = 0.65) for the probability of anemia.

Given decreasing marginal returns to inputs in the reduction of anemia, a natural question arising from these results is if substitution between incentives and resources is due to (i) the biological relationship between inputs and anemia (i.e., although more inputs are used, there is no effect on anemia because of biological constraints) or (ii) additional resources mitigating the effect of incentives on input use (i.e., anticipating decreasing marginal returns, administrators substituted resources and effort away from anemia reduction – as in Equation 17 of the model in Section 2). Our results for input use in Section 5.2 are consistent with the latter interpretation (we find direct evidence of substitution in input use). (We also note that efficacy trials of iron supplementation suggest that much larger reductions in anemia are biologically possible (Gera et al. 2007)).

5.2. Behavioral Responses Underlying Changes in Anemia

We next examine the underlying behavioral responses to our interventions that may have produced the changes in anemia described in Section 5.1. To do so, we focus on actions taken by administrators and subsequent responses among students and their parents – specifically, student consumption of iron-rich foods, direct iron supplementation, communication between parents and schools about anemia and its nutritional basis. For each family of outcome variables, we examine indices as described in Section 4.5.

Behavioral Responses Underlying Result 1: Large Incentives. We first consider the behavioral responses underlying Result 1 – that in the presence of the small block grant, the large incentive significantly reduced the probability of student anemia. Table 3 shows that the large incentive led administrators to increase vitamin supplementation and the provision of iron-rich foods by 0.14 standard deviations (sd) (Column 1, Row 2; adjusted p-value 0.104). This increase in iron-rich foods seems driven largely by home consumption (Row 2, Column 5, 0.187 sd, adjusted p-value 0.090).[46]

An interesting issue is if the increase is vitamin supplementation and provision of iron-rich foods occurred because school administrators with large incentives spent the block grant differently – or instead because they exerted more effort. As Figure 4 shows, reported use of block grants for different types of nutrition interventions (vitamins, food, fortification), and other uses is similar for incentive and non-incentives schools receiving a small grant, suggesting that greater anemia reduction due to incentives is driven by effort rather than differential allocation of the block grant.

In exploring how administrators were able to increase child consumption of iron rich foods at home, we examine contact with parents. Row 2, Column 8 of Table 3 reports a positive (but insignificant) increase in school contact of 0.130 standard deviations. However, Appendix Table 7 shows that estimates for several components of this index appear meaningful and important, albeit insignificant at conventional levels using adjusted p-values (largely because of the large number of hypotheses being tested ($11 \times 5 = 55$).[47] These results are suggestive that the large incentive led administrators to

[46] Sub-indices for supplements and food (including separate indices for food at home and school) were not explicitly specified in the pre-analysis plan.

[47] Specifically, the number of individual meetings between administrators and households over the past semester increased by 0.52 (Column 2, Row 2) – an increase of 59%; whether or not schools contacted

engage more regularly with households – specifically about nutrition and anemia – which in turn appears to have improved children's diets at home.

The finding that administrators responded to large incentives by engaging with households is important for at least two related reasons. First, it demonstrates innovation and the use of local knowledge in response to performance incentives that reward outputs (health outcomes) as opposed to those that rigidly reward the use of pre-specified inputs (such as vitamin consumption at school), as most performance incentives in the health sector do (Miller and Singer Babiarz, 2014). Second, for outcomes jointly produced with beneficiary households (like good child nutrition), it demonstrates the potential of performance incentives that reward outputs to minimize offsetting compensatory behavior among beneficiaries (a common finding among studies of school lunch programs, for example) (Jacoby 2002; Leonard 2003; Das et al. 2013).[48]

Behavioral Responses Underlying Result 2: Small Incentives. Second, we study the behavioral responses underlying Result 2 – that the small incentive did not reduce anemia prevalence. Table 3 (Column 1) shows that administrators with small incentives did not significantly increase the provision of supplements or Food (Row 1, Columns 1 to 3), nor did they increase their contact with households (Column 8) (Appendix Table 7,

---

parents about nutrition in the past semester rose by 12 percentage points (Column 3, Row 2) – an increase of 29%; and whether or not schools contacted parents about feeding children iron-rich foods rich in the past semester rose by 10 percentage points (Column 4, Row 2) – an increase of 47%. Note that the number of school-wide parent meetings and number of individual meetings with parents were not pre-specified to be part of this index.

[48] We speculate that the bureaucratic environment is one reason that administrators chose to work through households. Administrators may have viewed this strategy as a way to reduce anemia (and increase rewards) while avoiding the risk of career harm due to possible adverse events. This career harm may also be more severe under incentives if incentives altered perceptions of administrators' motivation for reducing anemia (analogous to how incentives may crowd-out effort if they alter the motives for prosocial tasks perceived by others – Bénabou and Tirole (2006)).

Row 1 also shows that none of the individual components of this index are statistically significant (even using unadjusted p-values).

Behavioral Responses Underlying Result 3: Large Block Grants. Third, we examine behavioral responses to large block grants, which reduced the prevalence of student anemia. The large block grant increased the provision of supplements and food by 0.22 sd (Table 3, Row 3, Column 1; adjusted p-value 0.006). This increase appears due to increases in both iron supplements (Column 2, 0.24 sd, adjusted p-value 0.051) and iron-rich food (Column 3, 0.196 sd, adjusted p-value 0.062).

Interestingly, the large block grant may have also increased school contact with parents – suggesting that administrators worked through households to reduce anemia without any explicit incentives to do so. Although the estimate for the index in Table 3 is not statistically significant (Row 3, Column 8), some estimates for index components are larger than those for incentives. This may reflect intrinsic or prosocial motivation – or a sense of obligation or organizational mission (Ashraf, Bandiera, and Jack, 2014). Furthermore, although the large block grant increased communication with households, the impact of block grants on food consumption at home is insignificant. We speculate that this could reflect less effort (relative to administrators with incentives) devoted to mitigating compensatory behavior by households in response to greater food provision at school (which seems to have increased, although not significantly, with large grants).

Behavioral Responses Underlying Result 4: Substitution between Large Incentives and Large Block Grants. Finally, with the combination of large incentives and large block grants, we find direct evidence of crowding-out of inputs consistent with our anemia estimates in Table 2 and with our model (Equation 17). Specifically, Table 3 shows that for vitamin supplementation and consumption of iron-rich foods (both at school and at home), estimates for the interaction between the large incentive and large block grant are negative, implying substitution (Row 5). The interaction between the small incentive and large grant is also negative, but not statistically significant. Overall, there is no evidence that resources and incentives are complements – and that at sufficiently high levels, they are substitutes.

5.3. Cost-Effectiveness

Finally, we examine the cost-effectiveness of each of our intervention combinations. In doing so, we consider both the sub-sample of children anemic at baseline and our full sample of children, and we present both "programmatic" cost-effectiveness (direct monetary program costs to the implementing organization) and social cost-effectiveness calculations. We calculate total social costs as the sum of: (a) programmatic costs; (b) the cost of public funds; and (c) costs incurred by households in responding to the interventions. From social costs we exclude incentive payments (apart from their contribution the cost of public funds), considering these payments to be transfers (Kremer, Miguel and Thornton 2004). Incentive payments may also not be considered a cost, but rather simply another way of allocating salary expenditures (Muralidharan and Sundararaman 2011).

Appendix Table 9 presents these results.[49] The key finding that we highlight is that although large block grants were as effective in reducing student anemia as large incentives, they were more expensive. First, considering full social costs and using the full sample, the cost per case of anemia averted was 1,453 yuan (about $227) in the large incentive/small block grant group – but 44% larger in the large block grant group (2,099

[49] See table notes for further details about these calculations.

yuan, or about $328). Second, the cost-effectiveness of these two interventions relative to each other is similar when we restrict our calculations to children anemic at baseline (as we do in Sections 4 and 5, following our pre-analysis plan). Specifically, the large incentive/small block grant intervention is approximately 50% more cost effective than large block grant intervention without incentives (723 yuan, or $113, per case of anemia averted vs. 1,447 yuan, or $226). Finally, considering calculating only programmatic costs and using children anemic at baseline, the cost-effectiveness of the large incentive/small block grant intervention is roughly one third of that of the large block grant (114 yuan, or about $18 vs 331 yuan, or about $52).[50]

6. Conclusion

This paper provides new evidence on how public sector managers respond to the provision of performance incentives. To the best of our knowledge, it is the first study to analyze how behavioral responses to performance pay interact with exogenously assigned levels of resources – a critical issue in the design of incentive systems under stringent resource constraints (as is common in many developing countries).

We report four key findings. First, when school administrators have fewer budgetary resources available to them, large performance incentives (with realized payments equivalent to a couple of months of annual salary) lead to substantial improvement in service delivery. This seems driven by greater effort rather than changes in budgetary resource allocation. In particular, school administrators demonstrated the

[50] Although the superior cost-effectiveness of large incentives over larger resources strictly applies only to the specific incentive/block grant contracts that we implemented, similar results are reported by others comparing incentive contracts with in-kind input transfers (Lavy 2002, Muralidharan and Sundararaman, 2011).

ability to innovate, working through their students' parents to alter nutritional practices at home. Second, smaller incentives (one tenth the size of the larger ones) were ineffective on average and had negative effects on pro-socially motivated administrators. Third, even absent explicit performance incentives, increasing school administrators' budgets led to important improvements in performance (but was considerably less cost-effective than using performance incentives), implying the presence of other motives – potentially including intrinsic ones – in our context.

Fourth, we find that performance incentives and unrestricted grants do not interact when incentives are small – and importantly, are substitutes in the production of health when incentives are large. The degree of substitution is substantial: at the policy-relevant levels that we study, increasing the size of unrestricted block grants completely crowds-out the effect of incentives (and vice-versa). This finding is consistent with our model's predictions that the effect of incentives will be small when budgetary resources are large because the marginal productivity of anemia-reducing inputs approaches zero – and our empirical results suggest that this happens quite quickly. This is an important result for resource-poor environments in which both budgetary resources and performance incentives are used simultaneously as policy levers for improving the quality of public service delivery.

Overall, among public sector administrators in rural China, we find evidence that appropriately designed performance incentives (sufficiently large, and absent substantial discretionary resources) can improve public sector service delivery – and ultimately, child outcomes. Despite the bureaucratic environment, our study suggests that performance pay can be an effective approach to motivating public sector managers.

References

Alcázar, Lorena, F. Halsey Rogers, Nazmul Chaudhury, Jeffrey Hammer, Michael Kremer, and Karthik Muralidharan. 2006. "Why Are Teachers Absent? Probing Service Delivery in Peruvian Primary Schools." International Journal of Educational Research 45 (3): 117–36.

Alderman, Harold, and Donald Bundy. 2012. "School Feeding Programs and Development: Are We Framing the Question Correctly?" The World Bank Research Observer 27 (2): 204–21.

Allen, Lindsay H., Bruno De Benoist, Omar Dary, and Richard Hurrell. 2006. "Guidelines on Food Fortification with Micronutrients." Geneva: World Health Organization.

Anderson, Michael L. 2008. "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): 1481–95.

Ashraf, Nava, Oriana Bandiera, and B. Kelsey Jack. 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery." Journal of Public Economics 120 (December): 1–17.

Athey, Susan, and John Roberts. 2001. "Organizational Design: Decision Rights and Incentive Contracts." The American Economic Review 91 (2): 200–205.

Balarajan, Yarlini, Usha Ramakrishnan, Emre Özaltin, Anuraj H Shankar, and SV Subramanian. 2011. "Anaemia in Low-Income and Middle-Income Countries." The Lancet 378 (9809): 2123–35.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2007. "Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment." The Quarterly Journal of Economics 122 (2): 729–73.

Bandiera, Oriana, Andrea Prat, and Tommaso Valletti. 2009. "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." The American Economic Review 99 (4): 1278–1308.

Banerjee, Abhijit, Angus Deaton, and Esther Duflo. 2004. "Health, Healthcare, and Economic Development: Wealth, Health, and Health Survices in Rural Rajisthan." The American Economic Review 94 (2): 326–30.

Banerjee, Abhijit, and Esther Duflo. 2006. "Addressing Absence." The Journal of Economic Perspectives 20 (1): 117–32.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Dhruva Kothari. 2010. "Improving Immunisation Coverage in Rural India: Clustered Randomised Controlled Evaluation of Immunisation Campaigns with and without Incentives." BMJ: British Medical Journal 340.

Basinga, Paulin, Paul Gertler, Agnes Binagwaho, Agnes Soucat, Jennifer Sturdy, and Christel MJ Vermeersch. 2011. "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): 1421–28.

Behrman, Jere, Susan Parker, Petra Todd, and Kenneth Wolpin. 2015. "Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools." Forthcoming, Journal of Political Economy.

Bénabou, R., Tirole, J., 2006. Incentives and Prosocial Behavior. American Economic Review 96, 1652–1678.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, and Victor Pouliquen. 2013. Turning a Shove into a Nudge? A “Labeled Cash Transfer” for Education. Working Paper 19227. National Bureau of Economic Research.

Berendes, Sima, Peter Heywood, Sandy Oliver, and Paul Garner. 2011. “Quality of Private and Public Ambulatory Health Care in Low and Middle Income Countries: Systematic Review of Comparative Studies.” PLoS Med 8 (4): e1000433.

Besley, Tim, and Maitreesh Ghatak. 2007. “Reforming Public Service Delivery.” Journal of African Economies 16 (1): 127.

Bhutta, Zulfiqar, Jai Das, Arjumand Rizvi, Michelle Gaffey, Neff Walker, Susan Horton, Patrick Webb, Anna Lartey, and Robert E Black. 2013. “Evidence-Based Interventions for Improvement of Maternal and Child Nutrition: What Can Be Done and at What Cost?” The Lancet 382 (9890): 452–77.

Bloom, Erik, Elizabeth King, Indu Bhushan, Michael Kremer, David Clingingsmith, Benjamin Loevinsohn, Rathavuth Hong, and J. Brad Schwartz. 2006. Contracting for Health: Evidence from Cambodia.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2011. Does Management Matter? Evidence from India. National Bureau of Economic Research.

Bó, Ernesto Dal, Frederico Finan, and Martín A. Rossi. 2013. “Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service.” The Quarterly Journal of Economics 128 (3): 1169–1218.

Bobonis, Gustavo, Edward Miguel, and Charu Puri-Sharma. 2006. “Anemia and School Participation.” Journal of Human Resources 41 (4): 692–721.

Bruhn, Miriam, and David McKenzie. 2009. “In Pursuit of Balance: Randomization in Practice in Development Field Experiments.” American Economic Journal: Applied Economics 1 (September): 200–232.

Bundy, Donald AP Sheldon Shaeffer, Matthew Jukes, Kathleen Beegle, Amaya Gillespie, Lesley Drake, Sueng-hee Francis Lee, et al. 2006. “School-Based Health and Nutrition Programs.” In Disease Control Priorities in Developing Countries, 30:75. New York: World Bank, World Health Organization, National Institutes of Health.

Bundy, Donald, and Helen Guyatt. 1996. “Schools for Health: Focus on Health, Education and the School-Age Child.” Parasitology Today 12 (8): 1–14.

Burrows, T.L., H. Truby, P.J. Morgan, R. Callister, P.S.W. Davies, and C.E. Collins. 2013. “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): 613–18.

Buttenheim, Alison M., Harold Alderman, and Jed Friedman. 2011. Impact Evaluation of School Feeding Programs in Lao Pdr. SSRN Scholarly Paper ID 1736865. Rochester, NY: Social Science Research Network.

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

Chaudhury, Nazmul, and Jeffrey S. Hammer. 2004. “Ghost Doctors: Absenteeism in Rural Bangladeshi Health Facilities.” The World Bank Economic Review 18 (3): 423–41.

Cohen, Jessica, and Pascaline Dupas. 2010. “Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment.” The Quarterly Journal of Economics 125 (1): 1–45.

Dasgupta, Amil, and Yianis Sarafidis. 2009. “Managers as Administrators: Reputation and Incentives.” Journal of Economic Behavior & Organization 70 (1–2): 155–63.

Das, Jishnu, Alaka Holla, Veena Das, Manoj Mohanan, Diana Tabak, and Brian Chan. 2012. “In Urban And Rural India, A Standardized Patient Study Showed Low Levels Of Provider Training And Huge Quality Gaps.” Health Affairs 31 (12): 2774–84.

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. “School Inputs, Household Substitution, and Test Scores.” American Economic Journal: Applied Economics 5 (2): 29–57.

Das, Jishnu, and Jeffrey Hammer. 2004. “Strained Mercy: Quality of Medical Care in Delhi.” Economic and Political Weekly, 951–61.

———. 2007. “Money for Nothing: The Dire Straits of Medical Practice in Delhi, India.” Journal of Development Economics 83 (1): 1–36.

Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard. 2008. “The Quality of Medical Advice in Low-Income Countries.” The Journal of Economic Perspectives 22 (2): 93–114.

Deci, Edward L., and Richard M. Ryan. 1985. Intrinsic Motivation and Self-Determination in Human Behavior. Springer Science & Business Media.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999. “The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies.” The Review of Economic Studies 66 (1): 199–217.

Dixit, Avinash. 2002. “Incentives and Organizations in the Public Sector: An Interpretative Review.” The Journal of Human Resources 37 (4): 696–727.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. “Incentives Work: Getting Teachers to Come to School.” The American Economic Review 102 (4): 1241–78.

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. “Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya.” American Economic Review 101 (6): 2350–90.

Du, Shufa, Fengying Zhai, Youfa Wang, and Barry M. Popkin. 2000. “Current Methods for Estimating Dietary Iron Bioavailability Do Not Work in China.” The Journal of Nutrition 130 (2): 193–98.

Education Law of the Peoples Republic of China. (Adopted at the Third Session of the Eighth National People's Congress on March 18, 1995, promulgated by Order No.45 of the President of the People's Republic of China on March 18, 1995 and effective as of September 1, 1995). http://www.china.org.cn/english/education/184669.htm#3

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

Fehr, Ernst, and Armin Falk. 2002. "Psychological Foundations of Incentives." European Economic Review 46 (4–5): 687–724.

Finch, Clement. 1994. "Regulators of Iron Balance in Humans." Blood 84 (6): 1697–1702.

Food Safety News. 2012. "School Lunch Suspected in Chinese Food Poisonings" (http://www.foodsafetynews.com/2012/04/school-lunch-suspected-in-chinese-food-poisonings/#.VW90HlwrLdQ, Accessed May 2015).

FORHEAD Working Group on Food Safety. 2014. "Food Safety in China: A Mapping of Problems, Governance and Research" (http://webarchive.ssrc.org/cehi/PDFs/Food-Safety-in-China-Web.pdf, Accessed May 2015).

Francois, Patrick. 2000. "'Public Service Motivation' as an Argument for Government Provision." Journal of Public Economics 78 (3): 275–99.

Francois, Patrick, and Michael Vlassopoulos. 2008. "Pro-Social Motivation and the Delivery of Social Services." CESifo Economic Studies 54 (1): 22–54.

Gera, T., Sachdev, H., Nestel, P., Sachdev, S.S., 2007. Effect of Iron Supplementation on Haemoglobin Response in Children: Systematic Review of Randomised Controlled Trials: Journal of Pediatric Gastroenterology and Nutrition 44, 468–486.

Gertler, Paul J., and Christel Vermeersch. 2012. "Using Performance Incentives to Improve Health Outcomes." Working Paper, June.

Gibbons, Robert, and Kevin J. Murphy. 1992. "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence." Journal of Political Economy 100 (3): 468–505.

Glewwe, Paul, Nauman Ilias, and Michael Kremer. 2010. "Teacher Incentives." American Economic Journal: Applied Economics 2 (3): 205–27.

Gneezy, U., S. Meier, and P. Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." Journal of Economic Perspectives.

Gneezy, Uri, and Aldo Rustichini. 2000. "Pay Enough or Don't Pay at All." The Quarterly Journal of Economics 115 (3): 791–810.

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

Hall, Brian J., and Jeffrey B. Liebman. 1998. "Are CEOs Really Paid Like Bureaucrats?" The Quarterly Journal of Economics 113 (3): 653–91.

Hall, Brian J., and Kevin J. Murphy. 2003. "The Trouble with Stock Options." Journal of Economic Perspectives 17 (3): 49–70.

Holmstrom, Bengt. 1982. "Managerial Incentive Problems: A Dynamic Perspective." In Essays in Economics and Management in Honor of Lars Wahlbeck. Heksinki: Swedish School of Economics.

Holmstrom, Bengt, and Joan Ricart i Costa. 1986. "Managerial Incentives and Capital Management." The Quarterly Journal of Economics 101 (4): 835–60.

Horton, S., and J. Ross. 2003. "The Economics of Iron Deficiency." Food Policy 28 (1): 51–75.

Islam, Mahnaz, and John Hoddinott. 2009. "Evidence of Intrahousehold Flypaper Effects from a Nutrition Intervention in Rural Guatemala." Economic Development and Cultural Change 57 (2): 215–38.

Jacoby, Hanan G. 2002. "Is There an Intrahousehold 'flypaper Effect'? Evidence from a School Feeding Programme." The Economic Journal 112 (476): 196–221.

Jensen, Michael C., and Kevin J. Murphy. 1990. "Performance Pay and Top-Management Incentives." Journal of Political Economy 98 (2): 225–64.

Jukes, Matthew C.H, Lesley J Drake, and Donald A.P Bundy. 2008. School Health, Nutrition and Education for All: Levelling the Playing Field. CABI.

Kamenica, Emir. 2012. "Behavioral Economics and Psychology of Incentives." Annual Review of Economics 4 (1): 427–52.

Karlan, Dean, John A. List, and Eldar Shafir. 2011. "Small Matches and Charitable Giving: Evidence from a Natural Field Experiment." Journal of Public Economics 95 (5–6): 344–50.

Kazianga, Harounan, De Walque, Damien, and Harold Alderman. 2009. "Educational and Health Impacts of Two School Feeding Schemes: Evidence from a Randomized Trial in Rural Burkina Faso." World Bank Policy Research Working Paper, No. 4976.

Kremer, Michael, Nazmul Chaudhury, F. Halsey Rogers, Karthik Muralidharan, and Jeffrey Hammer. 2005. "Teacher Absence in India: A Snapshot." Journal of the European Economic Association 3 (2-3): 658–67. doi:10.1162/jeea.2005.3.2-3.658.

Kremer, Michael, and Edward Miguel. 2007. "The Illusion of Sustainability." The Quarterly Journal of Economics 122 (3): 1007–65.

Kremer, M., Miguel, E., Thornton, R., 2004. Incentives to Learn. National Bureau of Economic Research Working Paper No. 10971.

Laffont, Jean-Jacques, and Jean Tirole. 1993. A Theory of Incentives in Procurement and Regulation. MIT Press, Cambridge, MA.

Lavy, Victor. 2002. "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement." Journal of Political Economy 110 (6): 1286–1317.

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

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

———. 2000. "Performance Pay and Productivity." The American Economic Review 90 (5): 1346–61.

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

Leonard, Kenneth L., and Melkiory C. Masatu. 2010. "Using the Hawthorne Effect to Examine the Gap between a Doctor's Best Possible Practice and Actual Performance." Journal of Development Economics 93 (2): 226–34.

Lewis, Maureen. 2006. "Governance and Corruption in Public Health Care Systems." Center for Global Development Working Paper 78 (January): 1–57.

Luo, Renfu, Xiaobing Wang, Linxiu Zhang, Chengfang Liu, Yaojiang Shi, Grant Miller, Scott Rozelle, Elaine Yu, and Reynaldo Martorell. 2011. "Alarmingly High Anemia Prevalence in Western China." Southeast Asian Journal of Tropical Medicine and Public Health 42 (5).

Magarey, Anthea, Jane Watson, Rebecca K. Golley, Tracy Burrows, Rachel Sutherland, Sarah A. Mcnaughton, Elizabeth Denney-Wilson, Karen Campbell, and Clare Collins. 2011. “Assessing Dietary Intake in Children and Adolescents: Considerations and Recommendations for Obesity Research.” International Journal of Pediatric Obesity 6 (1): 2–11.
Martorell, Reynaldo, Melany Ascencio, Luis Tacsan, Thelma Alfaro, Melissa F. Young, O. Yaw Addo, Omar Dary, and Rafael Flores-Ayala. 2015. “Effectiveness Evaluation of the Food Fortification Program of Costa Rica: Impact on Anemia Prevalence and Hemoglobin Concentrations in Women and Children.” The American Journal of Clinical Nutrition 101 (1): 210–17.
McPherson, R.Sue, Deanna M Hoelscher, Maria Alexander, Kelley S Scanlon, and Mary K Serdula. 2000. “Dietary Assessment Methods among School-Aged Children: Validity and Reliability.” Preventive Medicine 31 (2): S11–33.
Miller, Grant, and Kimberly Singer Babiarz. 2013. “Pay-for-Performance Incentives in Low- and Middle-Income Country Health Programs.” NBER Working Paper, April.
Miller, Grant, Renfu Luo, Linxiu Zhang, Sean Sylvia, Yaojiang Shi, Patricia Foo, Qiran Zhao, Reynaldo Martorell, Alexis Medina, and Scott Rozelle. 2012. “Effectiveness of Provider Incentives for Anaemia Reduction in Rural China: A Cluster Randomised Trial.” BMJ 345: e4809.
Miller, Grant, Sean Sylvia, Marcos Vera-Hernandez. 2015. “Multiple Tasks and Multiple Rewards: Experimental Evidence on Performance Incentives, Alignment and Complementarity from Chinese Schools.” Working paper.
Muralidharan, Karthik, and Venkatesh Sundararaman. 2011. “Teacher Performance Pay: Experimental Evidence from India.” Journal of Political Economy 119 (1): 39–77.
Murphy, Kevin J. 1999. “Chapter 38 Executive Compensation.” In Handbook of Labor Economics, Volume 3, Part B: 2485–2563. Elsevier.
Nath, Anusha. 2015. “Bureaucrats and Politicians: How Does Electoral Competition Affect Bureaucratic Performance?” Working paper 269, Institute of Economic Development, Boston University.
Nokes, Catherine, Claire van den Bosch, and Donald A.P. Bundy. 1998. “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.
Olken, Benjamin A., Junko Onishi, and Susan Wong. 2014. “Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia.” American Economic Journal: Applied Economics 6 (4): 1–34.
Orazem, P., P. Glewwe, and H. Patrinos. 2008. “The Benefits and Costs of Alternative Strategies to Improve Educational Outcomes.” Copenhagen Consensus 2008 Challeng Paper: Education.
Oxman, A.D., and A. Fretheim. 2008. An Overview of Research on the Effects of Results-Based Financing. Nasjonalt kunnskapssenter for helsetjenesten (Norwegian Knowledge Centre for the Health Services).
Oyer, Paul, and Scott Schaefer. 2005. “Why Do Some Firms Give Stock Options to All Employees?: An Empirical Examination of Alternative Theories.” Journal of Financial Economics 76 (1): 99–133.

Pasricha, Sant-Rayn, Hal Drakesmith, James Black, David Hipgrave, and Beverley-Ann Biggs. 2013. "Control of Iron Deficiency Anemia in Low- and Middle-Income Countries." Blood 121 (14): 2607–17.

Prendergast, Canice. 2002. "The Tenuous Trade-off between Risk and Incentives." Journal of Political Economy 110 (5): 1071–1102.

Rasul, Imran, and Daniel Rogger. 2013. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." Working Paper, University College London.

Rogger, Daniel. 2014. The Causes and Consequences of Political Interference in Bureaucratic Decision Making: Evidence from Nigeria. Mimeo.

Salanié, Bernard. 2005. "The Economics of Contracts: A Primer." MIT Press Books.

Singh, Prakarsh. 2011. Performance Pay and Information: Reducing Child Malnutrition in Urban Slums. MPRA Working Paper 29403. University Library of Munich, Germany.

Sylvia, S., Shi, Y., Xue, H., Tian, X., Wang, H., Liu, Q., Medina, A., Rozelle, S., 2015. Survey using incognito standardized patients shows poor quality care in China's rural clinics. Health Policy Plan. 30, 322–333.

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

Taras, Howard. 2005. "Nutrition and Student Performance at School." Journal of School Health 75 (6): 199–213.

Thomas, Duncan, Elizabeth Frankenberg, Jed Friedman, Jean-Pierre Habicht, Mohammed Hakimi, Nicholas Ingwersen, Nathan Jones, et al. 2006. "Causal Effect of Health on Labor Market Outcomes: Experimental Evidence." Working Paper, California Center for Population Research, UCLA.

Thornton, Rebecca L. 2008. "The Demand For, and Impact Of, Learning HIV Status." American Economic Review 98 (5): 1829–63.

Tonin, Mirco, and Michael Vlassopoulos. 2014. Are Public Sector Workers Different? Cross-European Evidence from Elderly Workers and Retirees.

Westfall, P. and Young, S. 1993. Resampling-Based Multiple Testing, New York: Wiley

Wilson, James Q., 1989. Bureaucracy: What Government Agencies Do and Why They Do It. New York: Basic Books.

World Bank. 2004. "World Development Report 2004: Making Services Work for Poor People." World Bank, Washington, DC.

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

Xinhua News Agency, "China Spends 47 Bln Yuan Improving Students' Nutrition" Jan 30, 2015, http://news.xinhuanet.com/english/china/2015-01/30/c_133959908.htm, Accessed June 4, 2015.

Xu, Longqi, Senhai Yu, Zexiao Jian, JialunYang, Changqiu Lai, Xianjun Zhang, and Changqian Zheng. 1995. "Soil-Transmitted Helminthiases: Nationwide Survey in China." Bulletin of the World Health Organization 73 (4): 507–13.

Yip, Ray. 2001. “Iron Deficiency and Anemia.” In Nutrition and Health in Developing Countries, edited by Richard D. Semba and Martin W. Bloem, 327–42. Humana Press.

Yip, Winnie, Timothy Powell-Jackson, Wen Chen, Min Hu, Eduardo Fe, Mu Hu, Weiyan Jian, Ming Lu, Wei Han, and William C. Hsiao. 2014. “Capitation Combined With Pay-For-Performance Improves Antibiotic Prescribing Practices In Rural China.” Health Affairs 33 (3): 502–10.

Figure 1: Study Regions

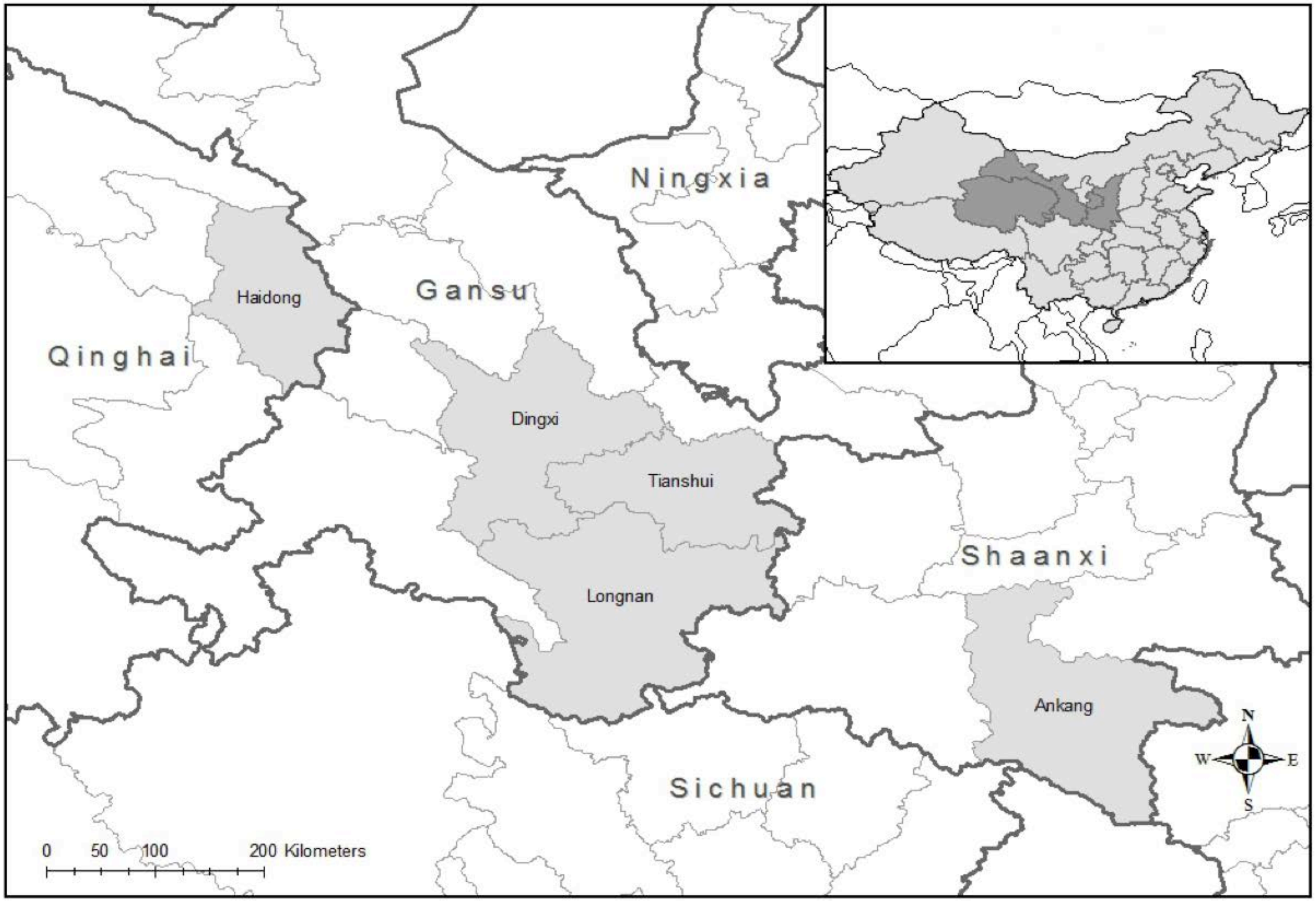


Figure 2: Experimental Design

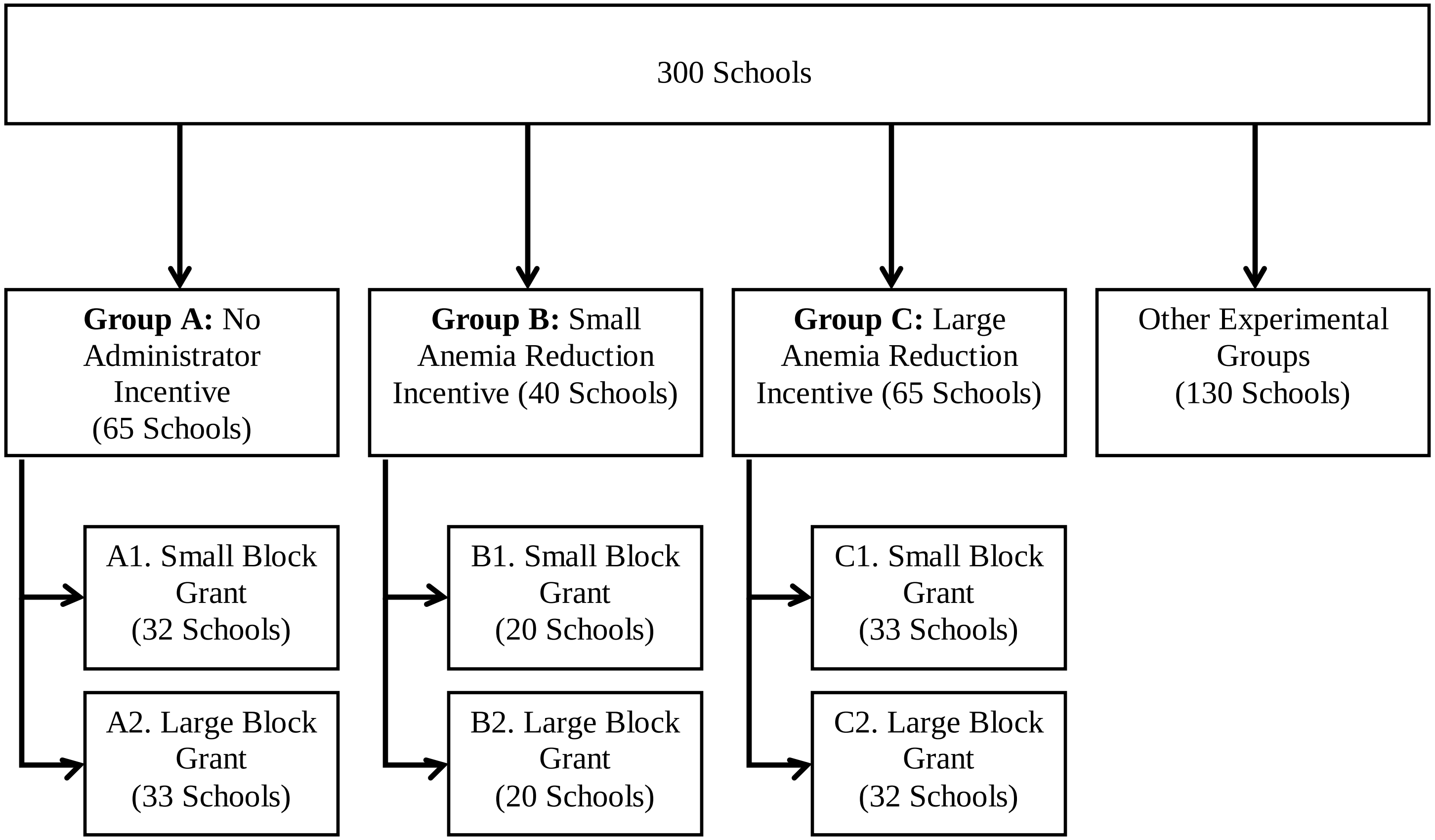


Figure 3: Data Collection and Intervention Timeline

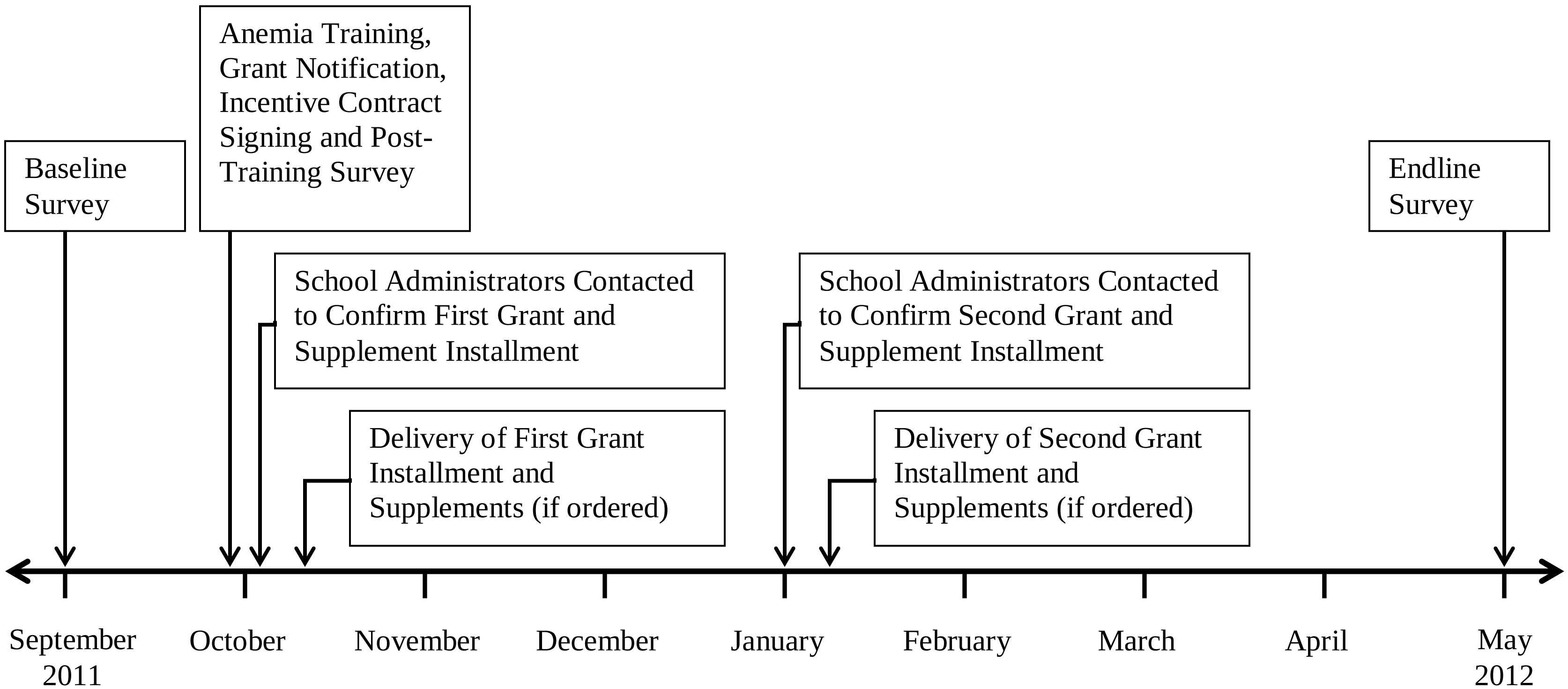
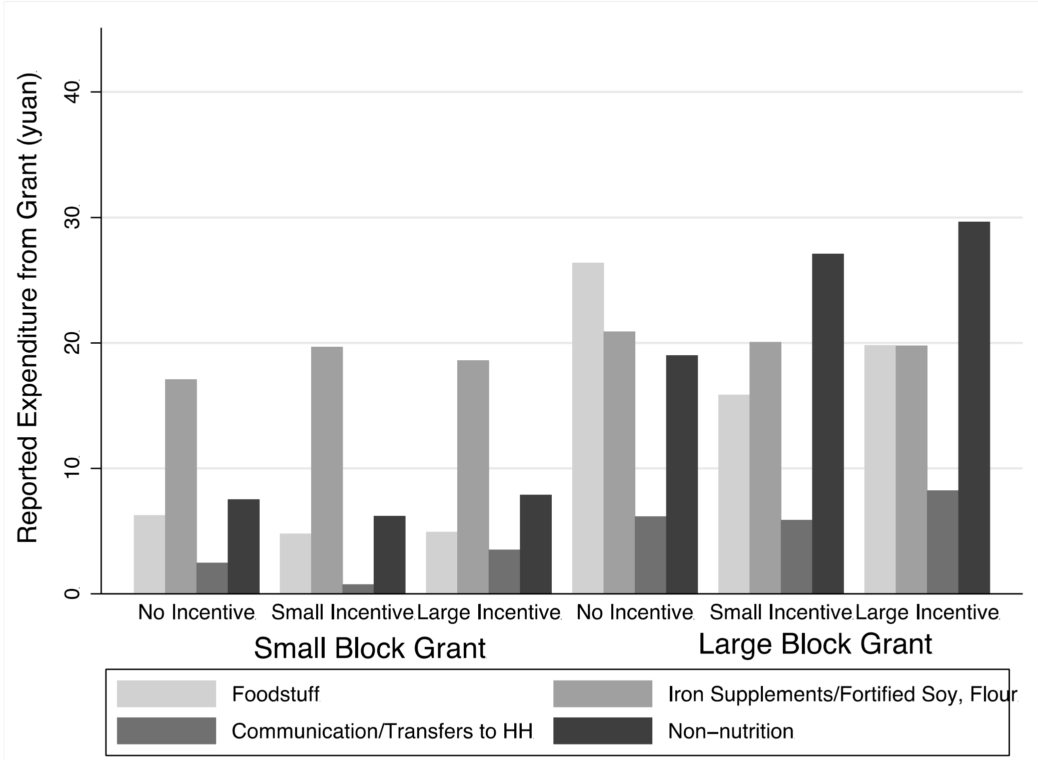


Figure 4: Reported Use of Block Grants by Category



NOTES: Figure shows mean values of reported use of block grants by experimental group from the endline survey. Expenditure amounts are per student.

Figure 5: Distributions of Hemoglobin Concentration

Figure 5A: Small Block Grant

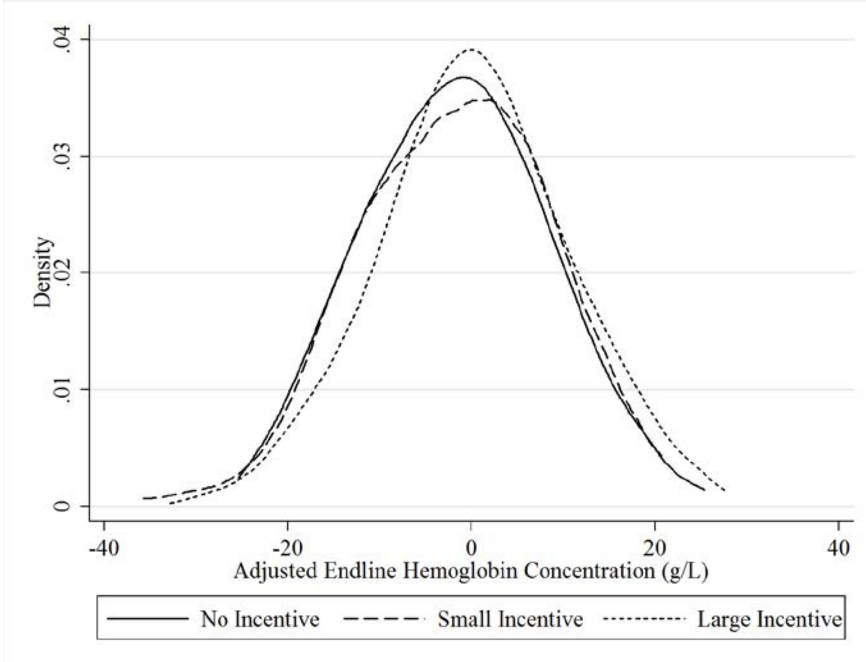
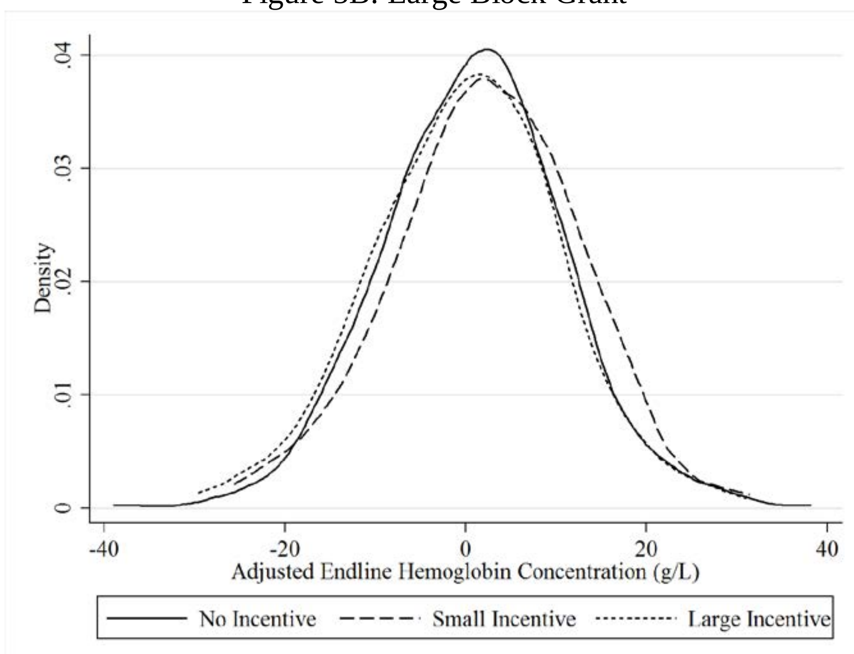


Figure 5B: Large Block Grant



NOTES: Figures plot the distributions of student level hemoglobin concentrations (in g/L) at endline across incentive treatment groups separately by small and large block grant groups. Endline Hb concentrations are adjusted for pre-specified baseline control variables. Kolmogrov-Smirnov P-values for Small Anemia Incentive vs. No Incentive are in 0.93 Panel A and 0.12 in Panel B. For Large Anemia vs. No Incentive these are in 0.02 Panel A and 0.24 in Panel B.

Table 1: Descriptive Statistics and Balance Check

	No Incentive, Small Grant Group		Coefficient (standard error) on:						
	Mean	SD	Small Incentive	Large Incentive	Large Grant	(Small-Incentive)X (Large Grant)	(Large Incentive)X (Large Grant)	Joint Test P-value: All Coefficients=0	Observations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Child Characteristics									
(1) 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)	0.420	2051
(2) 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)	0.914	2051
(3) 5th Grade (0/1)	0.468	--	0.055* (0.032)	-0.002 (0.029)	-0.003 (0.031)	-0.106** (0.050)	0.016 (0.042)	0.177	2051
(4) Female (0/1)	0.530	--	0.003 (0.044)	-0.021 (0.035)	-0.009 (0.039)	0.001 (0.060)	0.044 (0.052)	0.945	2051
(5) Times Consumed Meat in Past Week (incl. Chicken, Pork, Beef, Lamb)	3.922	4.145	-0.534 (0.618)	-1.293*** (0.453)	-0.352 (0.709)	-0.039 (0.888)	0.909 (0.790)	0.091*	2051
Panel B: School Characteristics									
(6) 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)	0.725	167
(7) Has Kitchen (0/1)	0.067	--	0.135 (0.099)	0.068 (0.077)	0.054 (0.085)	-0.071 (0.161)	-0.052 (0.120)	0.732	167
(8) 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)	0.185	167
(9) 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)	0.918	167
(10) 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)	0.985	167
Panel C. School Administrator Characteristics									
(11) Male (0/1)	0.967	--	-0.015 (0.051)	0.028 (0.032)	0.038 (0.034)	0.014 (0.053)	-0.070 (0.046)	0.606	167
(12) 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)	0.383	167
(13) Higher Education Degree (0/1)	0.900	--	0.018 (0.092)	-0.007 (0.081)	-0.107 (0.093)	0.032 (0.136)	-0.007 (0.126)	0.558	167
(14) 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)	0.137	167
(15) Monthly Base Salary (yuan)	1855.067	706.106	-57.049 (196.310)	-110.575 (178.286)	-36.880 (182.302)	-312.491 (305.716)	-35.944 (247.052)	0.602	167

NOTES. Data source: baseline survey. Table uses sample of children testing anemic at baseline. Children are considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines). The first and second columns show the mean and standard deviation in the comparison (small grant, no incentives) group. Columns 3 through 7 show coefficients and standard errors from a regression of each characteristic on indicators for incentive and large grant treatment group indicators and there interactions, controlling for randomization strata. Column 8 shows the p-value from a test that coefficients are jointly zero. All tests account for clustering at the school level. *, **, and *** indicate significance at 10%, 5% and 1%.

Table 2: Impacts of School Administrator Anemia Reduction Incentives and Block Grant Size on Student Hemoglobin Concentration and Anemia Prevalence

Dependent Variable:		Anemic at Endline (Hb<120 g/L)	Hemoglobin Concentration (g/L)
		(1)	(2)
Panel A: Impacts Relative to Comparison (No Incentive, Small Grant) Group			
(1)	Small Incentive	-0.012	-0.387
		(0.040)	(1.101)
		[0.771]	[0.726]
		{0.972}	{0.792}
(2)	Large Incentive	-0.138*	2.567
		(0.039)	(1.044)
		[0.001]	[0.015]
		{0.064}	{0.285}
(3)	Large Grant	-0.145**	4.205**
		(0.038)	(1.123)
		[<0.001]	[<0.001]
		{0.047}	{0.045}
(4)	(Small Incentive)X(Large Grant)	-0.042	1.445
		(0.056)	(1.541)
		[0.453]	[0.350]
		{0.888}	{0.664}
(5)	(Large Incentive)X(Large Grant)	0.196*	-4.580
		(0.058)	(1.586)
		[<0.001]	[0.004]
		{0.072}	{0.173}
(6)	Observations	1923	1923
(7)	Mean in No Incentive, Small Grant Group	0.364	129.901
Panel B: P-values of Additional Comparisons			
(8)	Large vs. Small Incentive	[0.002]	[0.014]
		{0.089}	{0.285}
(9)	Large Incentive vs. Large Grant	[0.854]	[0.169]
		{0.972}	{0.597}
(10)	Large Incentive vs. Large Incentive + Large Grant	[0.141]	[0.080]
		{0.650}	{0.511}

NOTES. Table uses sample of children testing anemic at baseline. Children are considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines). Rows 1-5 in Panel A show estimated coefficients for treatment group indicators and interactions obtained by estimating equation (12) (controlling 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). Standard errors are shown in parentheses, unadjusted p-values are shown in square brackets and p-values adjusted for multiple inference are shown in curly brackets. Adjusted p-values were constructed using the free step-down resampling method of Westfall and Young (1993) with 10,000 iterations. *, **, and *** indicate significance at 10%, 5% and 1% based on adjusted p-values. Panel B shows unadjusted and adjusted p-values from tests between coefficients.

Table 3: Child and Household Reported Receipt of Supplements and Iron-Rich Food

Dependent Variable:	Index: Supplements and Food	Sub-index: Supplements	Sub-index: Food	Sub-index: Food at School	Sub-index: Food at Home	Index: Information	Sub-index: Information to Students	Sub-index: Information to Households
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Small Incentive	0.069	0.138	-0.022	-0.083	0.072	0.027	0.062	0.045
	(0.053)	(0.084)	(0.049)	(0.061)	(0.061)	(0.075)	(0.126)	(0.084)
	[0.201]	[0.104]	[0.657]	[0.179]	[0.243]	[0.724]	[0.623]	[0.594]
	{0.336}	{0.234}	{0.750}	{0.475}	{0.403}	{0.935}	{0.911}	{0.690}
(2) Large Incentive	0.144	0.158	0.131	0.106	0.187*	0.070	0.111	0.130
	(0.052)	(0.080)	(0.048)	(0.060)	(0.064)	(0.072)	(0.101)	(0.084)
	[0.006]	[0.052]	[0.008]	[0.079]	[0.004]	[0.332]	[0.272]	[0.124]
	{0.104}	{0.234}	{0.101}	{0.426}	{0.090}	{0.800}	{0.796}	{0.430}
(3) Large Grant	0.221***	0.241*	0.196*	0.220	0.189	0.162	0.136	0.233
	(0.048)	(0.072)	(0.063)	(0.085)	(0.075)	(0.076)	(0.107)	(0.113)
	[0.000]	[0.001]	[0.002]	[0.010]	[0.013]	[0.034]	[0.207]	[0.041]
	{0.006}	{0.051}	{0.062}	{0.175}	{0.126}	{0.326}	{0.768}	{0.352}
(4) (Small Incentive)X(Large Grant)	-0.208	-0.269	-0.113	-0.108	-0.150	-0.037	0.190	-0.314
	(0.078)	(0.116)	(0.088)	(0.107)	(0.105)	(0.123)	(0.178)	(0.148)
	[0.008]	[0.021]	[0.201]	[0.314]	[0.157]	[0.766]	[0.288]	[0.036]
	{0.104}	{0.186}	{0.508}	{0.475}	{0.403}	{0.935}	{0.796}	{0.352}
(5) (Large Incentive)X(Large Grant)	-0.301**	-0.289	-0.314**	-0.362*	-0.292	-0.117	-0.039	-0.356
	(0.072)	(0.106)	(0.088)	(0.107)	(0.110)	(0.108)	(0.151)	(0.146)
	[0.000]	[0.007]	[0.000]	[0.001]	[0.009]	[0.282]	[0.797]	[0.016]
	{0.011}	{0.124}	{0.030}	{0.050}	{0.122}	{0.800}	{0.911}	{0.271}
(6) Observations	1932	1932	1928	1925	1927	1932	1928	1548
(7) Mean in No Incentive, Small Grant Group	-0.044	-0.055	-0.048	-0.043	-0.053	-0.040	-0.017	-0.082

NOTES. Table uses sample of children testing anemic at baseline. Children are considered anemic if they have an altitude-adjusted hemoglobin concentration below 120 g/L (per WHO guidelines). Rows 1-5 show estimated coefficients for treatment group indicators and interactions obtained by estimating equation (12) (controlling for the baseline value of the dependent 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). The dependent variable in each regression is a summary index constructed using the GLS weighting procedure in Anderson (2008). Estimates for the individual components of each index are shown in Appendix Tables 5 and 6. Standard errors are shown in parentheses, unadjusted p-values are shown in square brackets and p-values adjusted for multiple inference are shown in curly brackets. Adjusted p-values were constructed using the free step-down resampling method of Westfall and Young (1993) with 10,000 iterations. *, **, and *** indicate significance at 10%, 5% and 1%. based on adjusted p-values.