Paying for Performance and Cost Effectiveness of Strategies to Combat Anemia in China

Pre-Analysis Plan

Grant Miller
Marcos Vera-Hernandez
Linxiu Zhang
Renfu Luo
Yaojiang Shi
Sean Sylvia
Alexis Medina
Scott Rozelle

This document outlines the analysis plan for analyzing the impacts of the "Paying for Performance and Cost Effectiveness of Strategies to Combat Anemia in China" trial. The goal of this document is to pre-specify our planned methodology and specifications for analyzing the impacts of the program.[1] In addition to the ex-ante planned analysis specified here, we may conduct additional exploratory analyses as these are nevertheless an important means to discovery. However, this document serves as a record of our planned analyses and all analyses not planned ex-ante will be marked as such when reporting results.[2]

1. Study Overview

The "Paying for Performance and Cost Effectiveness of Strategies to Combat Anemia in China" study is a randomized controlled trial implemented northwest China designed to test outcome-based performance pay based on school-level anemia prevalence given to school principals as a means of improving the effectiveness of school-based anemia reduction programs. The experiment involves more than 300 primary schools across Qinghai, Gansu and Shaanxi provinces (see map of study sites). Based on results from a pilot study (reported in Miller et al. (2012)), the study was expanded to further test how performance pay based on anemia prevalence compares to more traditional rewards based on student exam scores, how the two types of rewards affect student achievement, and the complementarity of the two types of rewards in their effects on anemia rates and academic performance. Further, an additional arm was added to the study to test how results varied by the size of incentives provided.

[1] This plan was written before endline data were entered and cleaned.
[2] Note that unforeseen factors may require adjustments to the plan specified in this document. In the event that adjustments are made, these will be noted including a justification for the adjustment.

Experimental Design

The experiment was designed as a cluster randomized trial with schools randomly allocated to one of five treatment arms:

ARM 1: Budget Support and Information Only (65 Schools)

Schools in this arm received an unconditional block grant ("budget support") of either 0.3 yuan or 0.7 yuan (USD 0.05 or 0.11) per child per school day (see below). The use of this grant was at the school principal's discretion – principals could use these funds to implement an anemia reduction program or could allocate these to other school functions. Funds were under the school principals' control over the course of the project and their use was not monitored.

In addition to budget support, school principals were given training on causes of anemia (mainly iron deficiency in China), known effective strategies of preventing and treating anemia (consumption of iron-rich foods and/or supplements), and the relationship between anemia and academic performance.

This arm serves as the "standard of care" control arm in the study. All schools in the study received these two interventions.

ARM 2: Large Anemia Reduction Incentive (65 Schools)

Principals in this arm received the same two interventions as ARM 1 and additionally were given an incentive contract stipulating rewards based on the net reduction in students with anemia between the beginning and end of the school year. The incentive contract was structured as

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

where N_b is the number of students found to be anemic at baseline and N_e is the number of these same students who were anemic at the time of the endline survey. The slope of the contract (125 yuan (RMB) per student reduction– approx. $20) was set according to expected performance (based on a pilot study in 2010) yielding approximately two month's base salary for the average principal.

ARM 3: Small Anemia Reduction Incentive (40 Schools)

The treatment in this arm was identical to ARM 2. The slope of the incentives contract in this arm, however, was reduced to 12.5 yuan per student (10% of the contract in ARM 2).

ARM 4: Test Score Incentive

Principals in this arm received the same two interventions as ARM 1 and were given an incentive contract stipulating rewards based on the average gain in student scores on standardized exams over the course of the school year. For every 1 point gain in the average of student scores on exams in math and Chinese (combined) principals received 800 yuan (approx. $125). Principals were not penalized for a reduction in average exams scores.

ARM 5: Dual Incentive

Principals in this arm received the same two interventions as ARM 1 and were given both the ARM 2 incentive contract based on reductions in anemia prevalence and the ARM 4 incentive contract based on standardized test score gains.

Within each arm each arm half of the schools were randomly selected to receive budget support of 0.3 yuan per student per school day and the other half were given 0.7 yuan per student.

Subsidy Amounts

Within each arm half of the schools were randomly selected to receive budget support of 0.3 yuan per student per school day and the other half were given 0.7 yuan per student.

Administration

The treatments described above were administered in October 2011. Principals were invited to information sessions where they received information on anemia, including strategies to reduce anemia, and strategies to raise test scores. After information sessions, principals were informed of the amount of the subsidy that they would receive and were given randomly assigned incentive contracts.

Power Calculations

Power calculations were conducted before the beginning of the trial by Montecarlo simulation using parameter estimates from the 2010 pilot study (Miller et al. 2012) using the subsample of anemic students. From the pilot study data, the intraclass correlation was estimated adjusting for a number of covariates (Lag value of the outcome variable, Number of Students in Schools, Whether the school has a canteen, Student-teacher ratio, Distance to the furthest village served by the school, Percent boarding students, county dummies). After adjustment, the intraclass correlation at the school level for hemoglobin was found to be 0.11.

In addition, the following parameters were assumed:

- 15% loss to follow-up
- 50 children tested per school
- 13 anemic at baseline tested at follow-up accounting for loss

- 30.6% anemia rate (at 120 g/L cutoff)

Monte Carlo simulations using 500 replications were run on the regression

$$Hb_{is} = \alpha + \beta_1 Test\ Incentive_s + \beta_2 Anemia\ Incentive_s + \beta_3 (Anemia\ Incentive_s) * (Test\ Incentive_s) + \varepsilon_{is}$$

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

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

Sampling

Areas in China's northwest known to have relatively high anemia rates were included in the sampling frame. Schools in a total of 36 counties in Qinghai, Gansu, and Shaanxi provinceswere assessed for inclusion in the study.

A canvass survey covering the 36 counties above was completed in August 2011. This survey yielded a list of all primary schools in these counties and sizes of student population for each as reported by the local education ministry.

From this list, schools were excluded that had less than 150 students total (since these school are not likely to have 50 or more 4th and 5th grade students) or more than 300 students. Using the final list of schools, 300 townships were randomly selected and one school was randomly selected from each township. Where available, backup schools were also selected.

Randomization

Following the baseline survey in September 2011, the 300 schools were randomly allocated to each of the five main experimental arms and between the two subsidy levels within each arm. To improve the power of the experiment, the randomization process was stratified on baseline mean school-level hemoglobin concentration and mean combined standardized math and Chinese exam scores.

Data Collection

The baseline survey for the trial was conducted in September of 2011 and the endline was conducted in May 2012. Both rounds collected detailed information on students, schools, school principals, teachers, and households. When principals were given contracts and block grants in early October 2011, they were also given an additional short survey that collected information on

their intended use of the budget subsidy (though it was made clear that they were free to change their minds over the course of the study). At the time of the baseline survey, a telephone survey was also conducted with the school district superintendents of all sample schools which collected information on school districts including incentives and monitoring of school principals.

This first set of primary outcome variables for the trial are altitude-adjusted student hemoglobin concentration (Hb) and anemia defined as Hb<120 g/L. To collect hemoglobin concentration measurements, nurses from Xi'an Jiaotong Medical School accompanied enumerators during the baseline and endline surveys. Hemoglobin levels were measured on-site (at schools) using HemoCue Hb 201+ systems – a procedure considered state-of-the-art.

The second set of outcomes are student scores grade-appropriate standardized exams in math and Chinese given to students during the baseline and endline surveys. One-half of the students in each class were randomly assigned to take either the math or Chinese exam. Questions for the math exam were taken from the question bank of the Trends in International Mathematics and Science Study (TIMSS) and the Chinese test was designed with the help of the Shaanxi Provincial Education Bureau. The exams were extensively pretested to ensure they adequately captured variation in student achievement. In all analysis exam scores will be normalized by the control group distribution.

We will also collect qualitative data to further analyze the effects of the trial. Qualitative data includes open-ended questions on strategies collected as part of the survey as well as structured interviews collected in separate school visits following the endline survey. This data will be used to provide anecdotal evidence and will also be coded and used in further quantitative analysis.

2. Estimating Equations

General Econometric Framework

Our general approach for estimating treatment effects of the interventions will be to regress outcomes related to each hypothesis measured at follow-up on dummy variables indicating treatment status, baseline value of variables that we specify below, school characteristics measured at baseline, county fixed effects and stratum fixed effects.

For children level outcomes (e.g. hemoglobin level, anemia status, and tests scores) we will use the following specification[3]:

$$y_{isct} = \alpha + \beta T_{sc} + \theta Y_{isct-1} + \gamma X_{zt-1} + \mu_c + s_s + u_{izt}, \; t=1 \qquad (1)$$

[3] This is an ANCOVA specification. McKenzie (2011) discusses gains in efficiency from using ANCOVA over a difference-in-difference estimator in estimating treatment effects from randomized experiments. In particular, the ratio of the DID variance to the ANCOVA variance is $2/(1+\rho)$ where ρ is the autocorrelation of the outcome variable.

where y_{isct} is the outcome of interest for child i in school s located at county c at time t (which is 1 as there is only one round of follow-up data), T_{sc} is a vector of treatment status dummies (dummies for incentive arms and dummies for size of budget support)[4], Y_{isct-1} is the baseline value of the dependent variable (hemoglobin concentration or exam score as appropriate); a vector of child characteristics at baseline including child's age, class-year, and gender, child's hemoglobin level at baseline, and child's test score at baseline; $X_{z,t-1}$is a vector of child characteristics at baseline (including child's age, class-year, and gender) and school characteristics included when estimating power and also a dummy variable indicating school participation in the National Nutrition Subsidy Program which started during the project); μ_c are county fixed effects; and $s_s$are fixed effects for strata into which schools were classified during randomization into treatments.[5] In order to test some hypothesis, we will replace μ_c for county level characteristics to test for heterogeneous effects for variables that vary at the county level (such as strength of monitoring exerted by the school inspector).

In parallel to the way that the sample size calculations were carried out, our main specification will be estimated using children who were anemic at baseline. To account for possible measurement error in hemoglobin measurements and to assess the impact of the interventions on students at greatest risk of being anemic at the time of the baseline survey, we will also estimate (1) using a sample of students that we predict to be anemic at baseline on the basis of observable characteristics. To predict these students we will estimate

$$y_{it} = \alpha + X'\beta + \mu_c + \varepsilon_i$$

by maximum likelihood using students in the control group. y_{it} is student anemia status (hb<120 g/L) at baseline; the vector X will contain observable characteristics at baseline including student age, student sex, baseline hemoglobin concentration, mother's and father's education, and dummy variables indicating household possession of a list of durable goods and housing characteristics; μ_c are county dummy variables. Based on these parameter estimates obtained using only the control group, we will predict student anemia status at baseline for the rest of the sample. Students predicted to be anemic will then be included in a second estimation sample. Analysis using this sample of students will be the same as the analysis using the sample of students anemic at baseline; however, because the sample will be estimated, standard errors will be calculated by bootstrapping the full procedure using 500 replications.

For comparison, we will also conduct estimates using the full sample of students (both anemic and not).

[4] Note that, depending on the focus of the analysis, estimations will be done using a subset of data corresponding the the experimental arms of interest.

[5]The randomization was stratified by school-level mean hemoglobin concentration and mean math and Chinese exam scores (measured at baseline). We will add strata indicators to the estimating equations. We did not stratify the randomization by county because doing so was complicated by our decision to select only one school per town.

The parameter of interest, average treatment effect, is given by β_1.To compare the relative effectiveness of each of the treatments we will test equality of the coefficients using Wald tests. In addition, one of our main hypotheses in the experiment (below) concerns the complementarity of anemia and test incentives. To test this we will test the hypothesis $H_0: (\beta_2 + \beta_4) - \beta_5 = 0$ where β_2 and β_4 are the coefficients on ARMS 2 & 4 and β_5 is the coefficient on ARM 5.

For household, teacher, school and principal level outcomes, we will use equation (1) with Y_{isct-1} being the dependent variable of the regression measured at baseline (when available). If the dependent variable of the regression is missing at baseline for more than 10% of the sample, we will also report results omitting from the set of controls the same dependent variable of the regression measured at baseline. For household level outcomes, the regression will be at the household level and variables will be averaged over children in cases where there is more than one child in the household.

Standard Errors

When the dependent variable is defined at the child, teacher or household level, inference will be conducted clustering the standard errors at the school level using the cluster-corrected Huber-White estimator. For dependent variables measured at the school and/or principal level, standard errors will be computed robust to heteroskedasticity (using the standard formula).

Distributional Analyses

The nature of the intervention and the incentives put in place imply that it is plausible that some parts of the distribution of the outcome variable were more affected than others. To test hypotheses dealing with the distribution of the outcome variable, we will use quantile regressions when the outcome variables are continuous.

We will consider three different specifications when estimating quantile regressions: (a) the specification analogous to (1), (b) the specification analogous to (1) but omitting $Y_{\text{isct-1}}$, and (c) specification with only T_{sc}. We estimate (a) for analogy to (1) and because of possibly large efficiency gains. However, we also estimate (b) because the interpretation of the quantile treatment effect conditional on baseline values of the outcome variable might not be particularly appealing, especially for policy makers. When we estimate quantile regressions, we will use the entire sample, not only that of anemic children.

We are not aware of analytic methods to compute standard errors that account for clustering using quantile regressions. Hence, inference will be carried out using block-bootstrap (having the school as the block) on regressions with students, teachers and household level dependent variables.

Subgroup Analyses and Interaction Effects

Some hypothesis will require to test whether the effect of any of the treatment arms is heterogenous according to some observable characteristic. In case this observable characteristic is binary, we will test for the presence of heterogenous effects by including the corresponding binary variable in regression (1) together with the multiplication of this binary variable with each of the treatment arm binary variables. If the observable characteristic is continuous, we will test for heterogenous effects in two different ways:

(1) Assuming linearity: we will introduce the continuous variable in the regression together with the interaction of this variable with each of the treatment arms
(2) Allowing for non-linear effects: we will create a new binary variable which takes value 1 if the continuous variable is above or equal to the median , and 0 otherwise. We will introduce in regression (1) this new binary variable together with each of the interactions between such variable and each treatment arm binary variable. We will apply the same for integer but non-binary variables. In those circumstances, it could happen that there is a mass point at the median in such a way that the new binary variable takes value 1 for 75% of the sample or more. In those circumstances, we will define the new binary variable as taking value 1 if the value of the continuous variable is strictly larger than the median, and 0 if it is equal to the median or smaller.

In addition we will test heterogeneous treatment effects by baseline hemoglobin status and test scores by estimating separate models for observations defined by quartiles of these baseline distributions.

As indicated above, our main specification will be estimated on (1) solely children who were anemic at baseline, or who were predicted to be anemic on the basis of observables by an econometric model that we will estimate.

The specific variables that will be used in subgroup analyses are listed below under "Hypotheses of Heterogeneity in Response."

Multiple Outcomes

For each hypothesis, we will report a mean index which combines the information of closely inter-related outcomes. We will compute this index as in Anderson 2008 (section 3.2.1) and report its associated p-value.[6] The individual outcomes that are part of the index (part of the same hypothesis) are defined in the "Hypothesis and Indicators" section below. For each individual outcome, we will also report the results for individual outcomes. In addition to normal

[6] The procedure suggested increases efficiency by ensuring that outcomes that are highly correlated with each other receive less weight.

p-values we will also report p-valuesadjusted for multiple testing within the hypothesis[7]so as to control the Familywise Error Rate Control (using the free step-down resampling method as in Westfall and Young 1993). For hypotheses regarding heterogeneity in response, we will treat each variable as a separate hypothesis (i.e. tests for heterogeneous effects will not be adjusted).

Alternative Specifications for Presentation to Medical Audiences

In addition to the primary estimation approach discussed above, we will conduct some analyses using methods more familiar to medical audiences. The main difference will be the treatment of the error term: the basic specifications will remain the same, but will be estimated in hierarchical models with random effects for counties and schools. In addition, estimated partial effects for binary outcomes will be presented as is common in the medical literature, using relative risks and odds ratios.

3. Hypotheses and Indicators[8]

[Direct Effects of Incentives]

H1: Incentives provided to school principals based on reductions in school-level anemia prevalence will lead to improvements in student iron status.

Study Arm Comparison(s):

- Large Anemia Incentive vs. Budget Support Only
- Small Anemia Incentive vs. Budget Support Only

Outcomes/Indicators:

- Student hemoglobin concentration
- Student anemia status (hemoglobin concentration<120 g/L)

H2: Larger anemia incentives will be more effective than smaller anemia incentives at improving children's iron status as measured by hemoglobin concentration.

Study Arm Comparison(s):

- Large Anemia Incentive vs. Small Anemia Incentive

Outcomes/Indicators:

- Student hemoglobin concentration

[7] Some hypotheses include indicators at multiple levels. So, for example, when we have indicators both at the student and school level within a hypothesis. In these instances, the adjusted p-value will be based on the number of indicators within the same level for that hypothesis.
[8] Note that, regardless of phrasing, all hypotheses will be tested as two-tailed hypotheses.

- Student anemia status (hemoglobin concentration<120 g/L)

H3: Incentives (combined with information and budget support) provided to school principals based on student performance on standardized exams in Math and Chinese will lead to improvements in student performance on standardized exams in these subjects.

Study Arm Comparison(s):

- Test Score Incentive vs. Budget Support Only
- Dual Incentive vs. Budget Support Only

Outcomes/Indicators:

- Student scores on standardized exam in Math, normalized by distribution in control group
- Student scores on standardized exam in Chinese, normalized by distribution in control group

H4: Test score incentives are complementary with incentives for anemia reduction.

Study Arm Comparison(s):

- Test Score Incentive vs. Dual Incentive
- Large Anemia Incentive vs. Dual Incentive
- Small Anemia Incentive vs. Dual Incentive

Outcomes/Indicators:

- Student hemoglobin concentration
- Student anemia status (hemoglobin concentration<120 g/L)
- Student scores on standardized exam in Math, normalized by distribution in control group
- Student scores on standardized exam in Chinese, normalized by distribution in control group

[Indirect Effects]

H5: Anemia incentives will improve student academic performance.

Study Arm Comparison(s):

- Large Anemia Incentive vs. Budget Support Only
- Small Anemia Incentive vs. Budget Support Only

Outcomes/Indicators:

- Student scores on standardized exam in Math, normalized by distribution in control group
- Student scores on standardized exam in Chinese, normalized by distribution in control group

H6: Test incentives (combined with information on anemia) alone will improve student iron status compared to no incentives at all, but less than anemia incentives.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Small Anemia Incentive
- Test Incentive vs. Large Anemia Incentive

Outcomes/Indicators:

- Student hemoglobin concentration
- Student anemia status (hemoglobin concentration<120 g/L)

H7: Anemia incentive schools will have smaller gains in student academic performance.

Study Arm Comparison(s):

- Large Anemia Incentive vs. Budget Support Only
- Small Anemia Incentive vs. Budget Support Only
- Large Anemia Incentive vs. Test Incentive
- Small Anemia Incentive vs. Test Incentive

Outcomes/Indicators:

- Student scores on standardized exam in Math, normalized by distribution in control group
- Student scores on standardized exam in Chinese, normalized by distribution in control group

H8: Test incentive schools will have smaller gains in student iron status.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Large Anemia Incentive
- Test Incentive vs. Small Anemia Incneitve

Outcomes/Indicators:

- Student hemoglobin concentration
- Student anemia status (hemoglobin concentration<120 g/L)

H9 :Principals provided with test score incentives based on student performance on Chinese and Math exams will devote less resources to other subjects.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only

Outcomes/Indicators:

- Student grade on last end of semester county-level exam in Science (Normalized by county)
- Student grade on last end of semester county-level exam in English (Normalized by county)
- Teacher reported: Percent of "secondary" class time spent on Chinese or Math last week (Teacher Form, Section 2, Q 17)
- Teacher reported: Number of classes given per week (secondary class) (Teacher Form, Section 2, Q 18)
- Teacher reported: Percent of class time devoted to self-study (secondary subject) (Teacher Form, Section 2, Q 22)
- Teacher reported: Homeworks assigned in last week (secondary subject) (Teacher Form, Section 2, Q 26)
- Student Reported: Times science teacher absent last month (Student Form, Section 3, Q 54-8)
- Teacher reported: Hours teacher spent last week teaching and preparing secondary classes (Teacher Form, Section 5, Q 15)

H10: Providing principals with anemia reduction incentives will lead to a reduction of foods with comparatively lower iron benefits.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only

Outcomes/Indicators:

- Student reported consumption of foods not classified as iron-rich or promoting iron absorption.[9]
- School expenditure in past week on foods not classified as iron-rich or promoting iron absorption (foods other than meat, tofu, green vegetables and peppers, fruit) (reported by cafeteria worker)) (School Form, Section 5)

[9] Iron-rich food items include food items with high levels of iron or promote iron absorption: meat, tofu, green vegetables and peppers, and fruit. All other foods included in the food frequency questionnaire are considered non-iron rich.

H11: The effects of Anemia reduction incentives are in the lower tail of the hemoglobin distribution, while the effect of test score incentives occur more broadly in the distribution of test scores (which we will test using quantile regressions)**.**

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Large Anemia Incentive vs. Budget Support Only

Outcomes/Indicators:

- Student hemoglobin concentration
- Student scores on standardized exam in Math, normalized by distribution in control group
- Student scores on standardized exam in Chinese, normalized by distribution in control group

Mechanism Hypotheses

MH1: Principals in anemia incentive schools will devote more school resources to addressing anemia.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- School-level:
    - Total school expenditures on anemia reduction inputs[10] from budget support provided (principal reported) (Principal Form, Section 11, Q 1A through 16A)
    - School expenditures on anemia reduction inputs from other sources (principal reported) (Principal Form, Section 11, Q 1B through 16B)
    - Principal gave students vitamins over past year (principal reported) (Principal Form, Section 8, Q 6-11)

[10] Expenditures on anemia reduction inputs include: spending on iron-rich food for students (meat, tofu, vegetables, fruit, beans, multivitamins, and iron-fortified soy sauces and flour); providing information to students about anemia and iron deficiency; food subsidies for households; doctor visits for students; costs related to parent meetings; incentives for teachers and students; and other anemia reduction related expenditures identified by principals.

- School anemia-related expenditures per student (school accountant) (School Form, Section 3)
- Time students served meat at school per week (cafeteria worker) (School Form, Section 5, Q 81-2)
- School expenditure in past week on anemia-related food items (meat, tofu, green vegetables and peppers, fruit) (cafeteria worker)) (School Form, Section 5, Q 82 through 109)
- Number of times anemia-related food items on school menu in past week (record-based)) (School Form, Section 7, all)
- Teacher reported: Distributed vitamins to students this year (mode of teacher responses for school) (Teacher Form, Section 6, Q 8-3)

MH2: Principals in anemia incentive schools will devote more personal time and effort to addressing anemia.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Principal-level
  - Fraction of total hours worked principal spent on managing student meals/cafeteria operations (Principal Form, Section 9, Q 17)
  - Principal absences from school in past semester (principal reported) (Principal Form, Section 9, Q 13-1)
  - Total hours principal worked last week (principal reports) (Principal Form, Section 9, Q 14 though 24)

MH3: Principals with anemia incentives will be more likely to take steps to increase the effort that teachers devote to improving student iron status.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive

- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Teacher/Class level
    - Homeroom teacher identifies nutrition as a factor influencing study (Teacher Form, Section 6, Q 2)
    - Homeroom teacher selects vitamins, meat, tofu, vegetables, or fruit as important to student nutrition (Teacher Form, Section 6, Q 3)
    - Homeroom teacher knows of anemia (Teacher Form, Section 6, Q 4)
    - Homeroom teacher correctly identifies iron deficiency as a cause of anemia (Teacher Form, Section 6, Q 5)
    - Homeroom teacher correctly identifies eating meat as a way to address anemia (Teacher Form, Section 6, Q 7)
    - Teacher reported: Times principal spoke about student nutrition with teacher this semester (Teacher Form, Section 6, Q 8-1)
    - Teacher reported: Times principal spoke about anemia with teacher this semester (Teacher Form, Section 6, Q 8-2)
    - Teacher reported: school provided teacher training about nutrition (Teacher Form, Section 1, Q 4-7)

MH4: Principals with anemia incentives will be more likely to provide anemia-related resources to students and households.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Household/Student Level
    - Household received vitamins to give to student (Household Form, Section 3, Q 15-3)

- Student Reported: School provided vitamins to students (Student Form, Section 2, Q34-1)
- Student Reported: frequency of vitamin distribution Student (Student Form, Section 2, Q34-3)
- Student Reported: School provided vitamins to take home over the weekend (Student Form, Section 2, Q34-8)
- Student Reported: Number of other kids in class that do not take vitamins (Student Form, Section 2, Q34-9)
- Student Reported: Days given vitamins in May (Student Form, Section 2, Q34-10)
- Student reported consumption of meat in past week (from school) (Student Food Frequency Questionnaire)
- Student reported consumption of meat in past week (from home) (Student Food Frequency Questionnaire)
- Student reported consumption of green leafy vegetables in past week (from school) (Student Food Frequency Questionnaire)
- Student reported consumption of green leafy vegetables in past week (from home) (Student Food Frequency Questionnaire)
- Student reported consumption of fruit in past week (from school) (Student Food Frequency Questionnaire)
- Student reported consumption of fruit in past week (from home) (Student Food Frequency Questionnaire)

MH5: Principals with anemia incentives will be more likely to provide anemia-related information to teachers, students, and households.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Principal/School level
    - Princicpal discussed nutrition at parent meeting (principal reported) (Principal Form, Section 2, Q 6-1)
    - Principal discussed nutrition at teacher meeting (principal reported) (Principal Form, Section 3, Q 9,10)
    - Principal had teachers talk to students about nutrition (principal reported) (Principal Form, Section 8, Q 6-12)
    - Principal had teachers talk to students' parents about nutrition (principal reported) (Principal Form, Section 8, Q 6-14)
    - Principal spoke with teachers about nutrition (Principal Form, Section 8, Q 6-16)
- Household level
    - School contacted parent about student nutrition this semester (Household Form, 1, 3-1)
    - Given parent was contact by school about student nutrition, told household to give children foods rich in iron (Household Form, 1, 3-2)
    - Parent knows of anemia (Household Form, 1, 4-1)
    - Parent correctly identifies foods that can prevent anemia (Household Form, 1, 4-2)
- Student level
    - Student Reported: School provided nutritional information to student (Student Form, Section 2, 28-1)
    - Student Reported: Times school spoke with students about nutrition(Student Form, Section 2, 28-2)
    - Student Reported: School told students to eat meat(Student Form, Section 2, 28-3)
    - Student reports knowing of anemia (Student Form, Section 2, 33-1)
    - Student Reported: School provided information to students about anemia (Student Form, Section 2, 33-2)
- Teacher level
    - Times homeroom teacher held parent meetings and discussed nutrition (Teacher Form, Section 7, 2 and 5-1)

MH6: Principals with anemia incentives and smaller subsidies will be more likely to try to encourage households to address anemia (in ways that are cheaper to the school) than principals with anemia incentives and principals with larger subsidies.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

- Budget Support (Small) vs. Budget Support (Large)

Outcomes/Indicators:

- Household level
    - School contacted parent about student nutrition this semester (Household Form, 1, 3-1)
    - Given parent was contact by school about student nutrition, told household to give children foods rich in iron (Household Form, 1, 3-2)
    - Parent knows of anemia (Household Form, 1, 4-1)
    - Parent correctly identifies foods that can prevent anemia (Household Form, 1, 4-2)
- Student
    - Student reported consumption of meat in past week (from home) (Student Food Frequency Questionnaire)
    - Student reported consumption of green leafy vegetables in past week (from home) (Student Food Frequency Questionnaire)
    - Student reported consumption of fruit in past week (from home) (Student Food Frequency Questionnaire)

MH7: With Anemia incentives, principals will pursue narrow iron delivery approaches over broader nutritional improvement relative to principals with no incentives. Principals with anemia incentives will also pursue more iron-focused approaches than principals with test score incentives.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- The ratio of student-reported concumption of iron-rich food items to items not classified as iron-rich foods(Student Food Frequency Questionnaire).

- The ratio of principal reported expenditures on vitamins to other expenses classified as anemia reduction inputs (Principal Form, Section 11, Q 1A through 16A).

MH8: Principals in test incentive schools will devote more school resources to raising test scores in math and Chinese.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Anemia Incentive (Large and Small)
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Anemia Incentive (Large and Small)

Outcomes/Indicators:

- School level
    - School expenditures on academic inputs[11] from budget support provided (principal reported) (Principal Form, Section 11, Q 14-28)
    - School expenditures on academic inputs from other sources (principal reported) (Principal Form, Section 11, Q 14-28)
    - School acadmic expenditures per student (school accountant) (School Form, Section3)
- Teacher level:
    - Teacher reported: Number of classes given per week (math and chinese) (Teacher Form, Section 2, Q 2)
    - Teacher reported: Number of self-study classes each week (math or chinese) (Teacher Form, Section 2, Q 10)

MH9: Principals in test incentive schools will devote more personal time and effort to raising test scores in math and Chinese.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Anemia Incentive (Large and Small)
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only

[11] Expenditures on academic inputs include spending on parent meetings, incentivives for teachers and students, tutoring for students, text books, stationary, computers, desks and chairs, overtime for teachers, teacher training, and other items identified by principals as belonging to this category.

- Dual Incentive vs. Anemia Incentive (Large and Small)

Outcomes/Indicators:

- Principal Level:
    - Fraction of total hours worked principal spent on managing student instruction/study (Principal Form, Section 9, Q16)
    - Principal absences from school in past semester (principal reported) (Principal Form, Section 9, Q13-1)
    - Total hours principal worked last week (principal reportes) (Principal Form, Section 9, Q14-24)

MH10: Principals with test score incentives will be more likely to take steps to increase the effort that teachers devote to improving student exam performance.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Anemia Incentive (Large and Small)
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Anemia Incentive (Large and Small)

Outcomes/Indicators:

- Principal level
    - Principal offered teachers rewards for improving student exam scores (principal reported) (Principal Form, Section 4, Q 18-17)
- Student level
    - Number of times student given math homework in past week (Student Form, Section 3, Q 41)
    - Times math teacher returned corrected homework week (Student Form, Section 3, Q 42)
    - Number of times student given chinese homework in past week week (Student Form, Section 3, Q 44)
    - Times Chinese teacher returned corrected homework week (Student Form, Section 3, Q 45)
    - Student Reported: Times math teacher absent last month week (Student Form, Section 3, Q 54-6)

    - Student Reported: Times chinese teacher absent last month week (Student Form, Section 3, Q 54-7)
- Teacher level
    - Teacher reported: Teacher received subsidy from school (Teacher Form, Section 1, Q 3)
    - Teacher reported: School offered teacher rewards related to student grades (Teacher Form, Section 1, Q 4-2,4-4)
    - Teacher reported: School provided teacher training (Teacher Form, Section 1, Q 4-5)
    - Teacher reported: Times school provided teacher training (Teacher Form, Section 1, Q 4-6)
    - Teacher reported: Time in self-study class spent teaching (math and chinese) (Teacher Form, Section 2, Q 11)
    - Teacher reported: Homeworks assigned in last week (math or chinese) (Teacher Form, Section 2, Q 12)
    - Teacher reported: Teacher provided outside tutoring to students last week (Teacher Form, Section 2, Q 28)
    - Teacher reported: Times teacher provided outside tutoring to students last week (Teacher Form, Section 2, Q 29)
    - Teacher reported: Teacher paid for overtime work this semester (Teacher Form, Section 2, Q 32-1)
    - Teacher reported: Amount teacher paid for overtime work this semester (Teacher Form, Section 2, Q 32-2)
    - Teacher reported: Times principal observed class this semester (Teacher Form, Section 2, Q 33-1)
    - Teacher reported: Teacher hours worked per week (Teacher Form, Section 5, Q 1 & 2)
    - Teacher reported: Days teacher missed class this semester (for any reason) (Teacher Form, Section 5, Q 3-1)
    - Teacher reported: School requested that teacher work overtime this semester (Teacher Form, Section 5, Q 4-1)
    - Teacher reported: Times teacher worked overtime this semester (Teacher Form 5, 5-1)
    - Teacher reported: Hours teacher worked overtime this semester (Teacher Form 5, 6-1)
    - Teacher reported: Hours teacher spent last week grading homework (Teacher Form 5, 5-8)
    - Teacher reported: Hours teacher spent last week preparing classes Teacher (Teacher Form 5, 5-9)
    - Teacher reported: Hours teacher spent last week teaching (Teacher Form 5, 5-10)

- Teacher reported: Hours teacher spent last week meeting with parents (Teacher Form 5, 5-11)
- Teacher reported: Hours teacher spent last week in everyday management activities (Teacher Form 5, 5-12)
- Teacher reported: Hours teacher spent last week teaching and preparing main classes (Teacher Form 5, 5-14) (math and chinese)
- Teacher reported: Hours teacher spent last week tutoring students outside of class (Teacher Form 5, 5-13)

MH11: Principals with test score incentives will be more likely to provide resources to students and households related to improving exam score performance.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Anemia Incentive (Large and Small)
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Anemia Incentive (Large and Small)

Outcomes/Indicators:

- Principal level
  - Principal discussed improving student's study outcomes at parent meeting (Principal Form, Section 3, Q 6-2)
- Student level
  - Student Reported: School provided tutoring to student (Student Form, Section 3, Q 54-4)
- Teacher/Class level
  - Teacher reported: Students offered rewards for good grades (Teacher Form, Section 2, Q 14-11 and 14-12)
  - Times homeroom teacher held parent meetings and discussed academics (Teacher Form, Section 7, 2 and 5-1)

MH12: Schools with incentives and less budget support are more likely to work through households (for reducing anemia and/or improving academic performance).

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only

- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Budget Support (Small) vs. Budget Support (Large)
- Test Incentive vs. Budget Support Only
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only

Outcomes/Indicators:

- Principal/School level
    - Parent meetings called by principal (principal reported) (Principal Form, Section 2, Q 5-1)
    - Percent of parents attending parent meetings this semester (principal reported) (Principal Form, Section 3, Q 7-1)
    - Times principal held teacher meetings in past 2 weeks (Principal Form, Section 3, Q 8)
- Teacher/Class level
    - Teacher reported: Times teacher spoke to students' parents last week (Teacher Form, Section 2, Q 34)
    - Teacher reported: Times teacher visited households last week (Teacher Form, Section 2, Q 35)
    - Teacher reported: Percent of parents that showed up for parent meetings this semester (Teacher Form, Section 7, Q 6-1)
- Household/Student level
    - Times member from household participated in parent meetings at school (Household Form, Section 1, Q 1-1,2-1)
    - Household received subsidy from school (Household Form, Section 3, Q 14-1, 14-2)
    - School expenditure on parent meetings (principal reported)

MH13: In test incentive schools, more effort will be devoted to improving test taking

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Anemia Incentive (Large and Small)
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Anemia Incentive (Large and Small)

Outcomes/Indicators:

- Teacher/class level
    - Teacher reported: Number of test preparation classes this semester (math or chinese)(Teacher Form, Section 2, 14-1)
    - Teacher reported: Number of test given this semester (including pratice) (Teacher Form, Section 2, 14-2)
    - Teacher reported: Number of test given last semester (including pratice) (Teacher Form, Section 2, 14-3)
    - Teacher reported: Times this semester reviewed test-taking skills with students (Teacher Form, Section 2, 14-4)

MH14: Incentive schools and schools with less budget support will be more technically efficient than budget support only schools.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Budget Support (Small) vs. Budget Support (Large)
- Test Incentive vs. Budget Support Only
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only

Outcomes/Indicators:

- Increase in student hemoglobin concentration per yuan of budget support
- Mean value added to student standardized exam scores in math and Chinese per yuan of budget support
- Increase in student hemoglobin concentration per yuan of total school expenditure
- Mean value added to student standardized exam scores in math and Chinese per uan of total school expenditure
- School allocative and technical efficiency estimated using data envelopment analysis

MH15: Students in Anemia and Test Incentive schools will have better intermediate academic outcomes

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Test Incentive vs. Budget Support Only
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only

Outcomes/Indicators:

- Student level
    - Times student absent from class in past week (Student Form, Section 3, Q 36)
    - Times student turned in math homework in past week (Student Form, Section 3, Q 40)
    - Times student turned in Chinese homework in past week (Student Form, Section 3, Q 43)
    - Student grade on last end of semester exam in math Student (Student Form, Section 3, Q 46 through 48)
    - Student grade on last end of semester exam in Chinese (Student Form, Section 3, Q 49 through 51)
    - Student grade on last end of semester exam in Science (Student Form, Section 3, Q 52 through 54)
    - Student grade on last end of semester exam in English (Student Form, Section 3, Q 54-1 through 54-3)
- Class level
    - Teacher reported: Percent of children in class actively participating (last week) (math or chinese) (Teacher From, Section 2, Q8)
    - Teacher reported: Percent of children in class not concentrating (last week) (math or chinese) (Teacher From, Section 2, Q9)

MH16: The general health of students in schools with an anemia incentive will improve

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Student level
    - Times student absent due to illness in past week (Student Form, Section 3, Q55)
    - Student had trouble concentrating in class in last week (Student Form, Section 3, Q38)
    - Self-reported student health (Student Form, Section 4, Q56 through 58-12)

MH17: Principals with incentives will be less likely to divert budget support to other school functions.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Test Incentive vs. Budget Support Only
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only

Outcomes/Indicators:

- Principal level
    - Principal assessment of whether other principals would divert funds from nutrition to other school functions (Principal Form, Section 6, 11-7, 11-8)

MH18: Incentive schools will be more likely to assess fees on households

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Test Incentive vs. Budget Support Only
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only

Outcomes/Indicators:

- Household level
    - Fees paid to school by household (household reported) (Household Form, Section 3,Q 8-1 through 13-1)
    - Fees for school meals paid by household to school (Household Form, Section 3,Q 12-1)
    - Fees for academics paid by household to school (Household Form, Section 3,see grouping)

MH19: Principals with anemia incentives will be more likely to retain information from aneamia training and seek out anemia related information.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Principal level
    - Principal score on anemia knowledge quiz (Principal Form, Section 8, 3 through 6-8)
    - Principal sought additional information on anemia after training (Principal Form, Section 8, 6-9)

MH20: Incentives will cause principals to believe tasks are more difficult.

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Test Incentive vs. Budget Support Only
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only

Outcomes/Indicators:

- Principal level
    - Principal subjective expectation of ability to improve student iron status (Principal Post-information Survey, Section 2, Q 1-3)
    - Principal subjective expectation of ability to improve student exam scores (Principal Post-information Survey, Section 2, Q 3-4)

MH21: Households in anemia incentive schools might provide more or less iron-rich food (or foods promoting iron absorbtion) at home than households in the control arm. The direction will depend on whether crowding-in (due to the information that schools provide to households) or crowding-out (due to the food and vitamins that the child might receive in the school) dominates

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Household level
    - Household expenditure in past week on iron rich food items (Household Form, Section 4,Q 17 through 24-1)
    - Household expenditure on food items in past week (Household Form, Section 4,Q 16)
    - Household food expenditure as a fraction of total household expenditure in past month (Household Form, Section 5,Q 24-1 through 24-6)
- Student level
    - Student reported consumption of meat in past week (from home) (Student Food Frequency Questionnaire)
    - Student reported consumption of green leafy vegetables in past week (from school) (Student Food Frequency Questionnaire)
    - Student reported consumption of green leafy vegetables in past week (from home) (Student Food Frequency Questionnaire)
    - Student reported consumption of fruit in past week (from school) (Student Food Frequency Questionnaire)

- Student reported consumption of fruit in past week (from home) (Student Food Frequency Questionnaire)
- Student reported consumption of beans/tofu (from school) (Student Food Frequency Questionnaire)
- Student reported consumption of beans/tofu (from home) (Student Food Frequency Questionnaire)

MH22: Households in anemia incentive schools might expend more or less in cafeteria than households in the control arm. The direction will depend on whether crowding-in (schools can increase cafeteria fees due to the information that schools provide to households) or crowding-out (due to the food and vitamins that the school itself chooses to provide) dominates

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Household expenditure on school meals, during past semester (Household form, Section 3, Q 12-1).
- Household expenditure on food last week (Household Form, Section 4, Q18, 20, 22)
- Student reported consumption of food at home (Food Frequency Questionnaire)

MH23: Households in schools with test score incentives schools might expend more or less on education than households in the control arm. The direction will depend on whether crowding-in (due to the information/motivation that schools provide to households) or crowding-out (due to the due to the academic inputs that the child might receive in the school) dominates.

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Anemia Incentive (Large and Small)
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Anemia Incentive (Large and Small)

Outcomes/Indicators:

- Household expenditure on books (Household form, Section 3, Q8-2)
- Household expenditure on tutoring (Household form, Section 3, Q9-1)
- Household expenditure on study materials (Household form, Section 3, Q13-1)

MH24: Adult members in households in schools with incentives for anemia reduction might work more or less than adult members in the control arm. The direction will depend on whether household consumption increases or decreases

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Hours worked in past week by student's mother (Household Form, Section 2, Q 4-6)
- Hours worked in past week by student's father (Household Form, Section 2, Q 4-5)
- Student's mother migrated for work (Student Form, Section 1, Q15)
- Student's father migrated for work (Student Form, Section 1, Q11)

MH25: Adult members in households in schools with test score incentives might work more or less than adult members in the control arm. The direction will depend on whether household consumption increases or decreases

Study Arm Comparison(s):

- Test Incentive vs. Budget Support Only
- Test Incentive vs. Anemia Incentive (Large and Small)
- Test Incentive vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Anemia Incentive (Large and Small)

Outcomes/Indicators:

- Hours worked in past week by student's mother (Household Form, Section 2, Q 4-6)
- Hours worked in past week by student's father (Household Form, Section 2, Q 4-5)
- Student's mother migrated for work (Student Form, Section 1, Q15)
- Student's father migrated for work (Student Form, Section 1, Q11)

MH26: Iron consumption will increase or decrease for other children in the households of children in anemia incentive schools

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Household level
  - Frequency student's sibling ate meat in past week (Household Form, Section 6, Q 30 through 34)
  - Frequency student's sibling ate fruit in past week (Household Form, Section 6, Q 37)
  - Frequency student's sibling ate vegetables in past week (Household Form, Section 6, Q 38-1)
  - Students sibling given multivitamins (Household Form, Section 6, Q 38-2)

MH27: Children in anemia incentive schools will be more likely to receive vitamins closer to the end of the semester (May).

Study Arm Comparison(s):

- Anemia Incentive (Large and Small) vs. Budget Support Only
- Anemia Incentive (Large and Small) vs. Test Incentive
- Anemia Incentive (Large and Small) vs. Dual Incentive
- Dual Incentive vs. Budget Support Only
- Dual Incentive vs. Test Incentive

Outcomes/Indicators:

- Student Reported: Days given vitamins in May (Student Form, Section 2, Q34-10)

- Student Reported: Days given vitamins in April (Student Form, Section 2, Q34-11)
- Student Reported: Days given vitamins in March (Student Form, Section 2, Q34-12)
- Student Reported: Days given vitamins in February (Student Form, Section 2, Q34-13)

Hypotheses of Heterogeneity in Response[12]

[Size of Budget Support] (All will be tested using an interaction between incentives treatments and subsidy size treatment)

- The effect of principal incentives on student outcomes will be larger when principals have more resources.
- Targeting attempts will increase with fewer resources.
- Principals with smaller subsidies will be more likely to pursue home-based approaches and do so more intensively (crowding-in will be more likely).
- Principals with smaller subsidies and incentives will be more likely to reallocate existing schools resources in incentivized direction.
- Schools with smaller subsidies will be more efficient

[School System Characteristics]

- Principal has pre-existing test score incentive (Principal Form, Section 4, Q 1 & 10)
- School system gives principals rewards for good evaluation ranking (Principal Form, Section 5, Q 25)
- School system evaluates schools (based on student grades) (Principal Form, Section 5, Q 29)
- Index of principal decision making autonomy (Principal Form, Section 9)
- Times someone from school district made monitoring visit to schools in past year (Superintendent Form, Section 1, Q6)
- Times school district had principal meetings last year (Superintendent Form, Section 1, Q7)
- School district awards principals based on school rankings (Superintendent Form, Section 1, Q13)

[School Characteristics]

- School has performance pay for teachers (Principal Form, Section 4, Q 1)
- Weighting of teacher evaluations on exam scores (Principal Form, Section 4)
- School Size (School Form, Section 1, Q 2)

[12] Note that each of the items in this section will be treated as a separate hypothesis.

- Number Percentage of Boarding Students (School Form, Section 1, Q3)
- Average student travel time to school (Studnt Form, Section 1, Q16)
- Distance of the furthest village served by School (School Form, Section 1, Q17)
- School has Kitchen (School Form, Section 1, Q30)
- Percent of teaching forces that is "High" or "Special" level (School Form, Section 1, Q11 & 12)
- Student to Teacher Ratio (School Form, Section 1, Q2,4)
- School Infrastructure Index (School Form, Section 1, Q 25 through 33)
- Total School Income in Previous Semester (School Form, Section 2, Q 45)
- Estimated School Cost of 1 hour of instruction (School Form, Section 2, Q 60)
- Cost of iron-rich food basket (meat, tofu, green vegetables and peppers, fruit) (school accountant) (School Form, Section 5, Q 82-109)
- School participated in Nutrition Improvement Program (NIP) (School Form (endline), Section 4, Q 1)
- School length of exposure to NIP (School Form (endline), Section 4, Q 2)
- Average Teacher Internal Motivation (based on scales) (Teacher Form, Section 4, Q 1-8)

[School Principal Characteristics]

- Principal baseline subjective expectation of influencing student iron status (Principal Post-Information Form, Section 1, Q 11 through 13)
- Principal baseline subjective expectation of influencing student exam scores(Principal Post-Information Form, Section 1, Q 14 through 16)
- Principal score on anemia knowledge quiz at baseline (immediately after information session) (Principal Post-Information Form, Section 1, Q 3 through 10)
- Gender (Principal Form, Section 1, Q 2)
- Age (Principal Form, Section 1, Q 3)
- Ethnicity (Principal Form, Section 1, Q 4)
- Education Level (Principal Form, Section 1, Q 5)
- Years of Teaching Experience (Principal Form, Section 1, Q6)
- Tenure as Principal (Principal Form, Section 1, Q9)
- Base Salary (Principal Form, Section 3, Q1)
- Total Household Income (Principal Form, Section 3, Q5)
- Total Income (Principal Form, Section 3, Q1 through 3)
- Percentage weighting of yearly evaluation on student scores (Principal Form, Section 4, Q17)
- Principal believes seniority influences promotion (likert scale) (Principal Form, Section 5, Q34)
- Principal believes school performance evaluation influences promotion (likert scale) (Principal Form, Section 5, Q35)

- Subjective probability of promotion in 1 year (Principal Form, Section 5, Q36)
- Subjective probability of promotion in 5 years (Principal Form, Section 4, Q37)
- Principal believes educators (principal and teachers) have the most responsibility to raise student performance (compared to upper levels or parents and students) (Principal Form, Section 6, Q2)
- Principal believes educators (principal and teachers) know best how to improve student health (compared to upper levels or parents and students) (Principal Form, Section 6, Q3)
- Principal believes educators (principal and teachers) have the most responsibility to improve student health (compared to upper levels or parents and students) (Principal Form, Section 6, Q4)
- Thinks parents more effective than school in improving student health (Principal Form, Section 6, Q5)
- Believes most people can be trusted (Principal Form, Section 6, Q6)
- Subjective probability that teachers will perform a task well when asked (Principal Form, Section 6, Q7)
- Subjective probability that kitchen staff will do a task well when asked (Principal Form, Section 6, Q8)
- Risk preference (Principal Form, Section 6, Q9)
- Subjective discount rate (Principal Form, Section 6, Q10 and 11)
- Prosocial Motivation Score (Principal Form, Section 7, Q1,3,5,7)
- Intrinsic Motivation Score (Principal Form, Section 7, Q 2,4,6,8)
- Internal Motivation Score (Intrinsic and Prosocial) (Principal Form, Section 7)
- Believes nutrition influences grades (selected from list) (Principal Form, Section 8, Q 2)

[Student Characteristics]

- Household size (Student Form, Section 1)
- Number of siblings(Student Form, Section 1)
- Times household member met with teacher or principal last semester(Household Form, Section 1, Q1,3)
- Student gender (Student Form, Section 1, Q 1)
- Ethnic Group (Student Form, Section 1, Q 2)
- Student boards at school (Student Form, Section 1, Q 5)
- Distance from home to school (Student Form, Section 1, Q 16)
- Parent migration status (Student Form, Section 1, Q 11, 15)
- Parent education (Student Form, Section 1, Q 9, 13)
- Student likes meat, tofu (Student Form, Section 2, Q 29,32)
- Household Asset Index (Student Form, Section 5, Q 59-67)

- Student Baseline Chinese and Math Scores Student
- Student Baseline Hemoglobin Concentration Student

References

Anderson, M. (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 Statitstical Association* 103: 484, 1481-1495

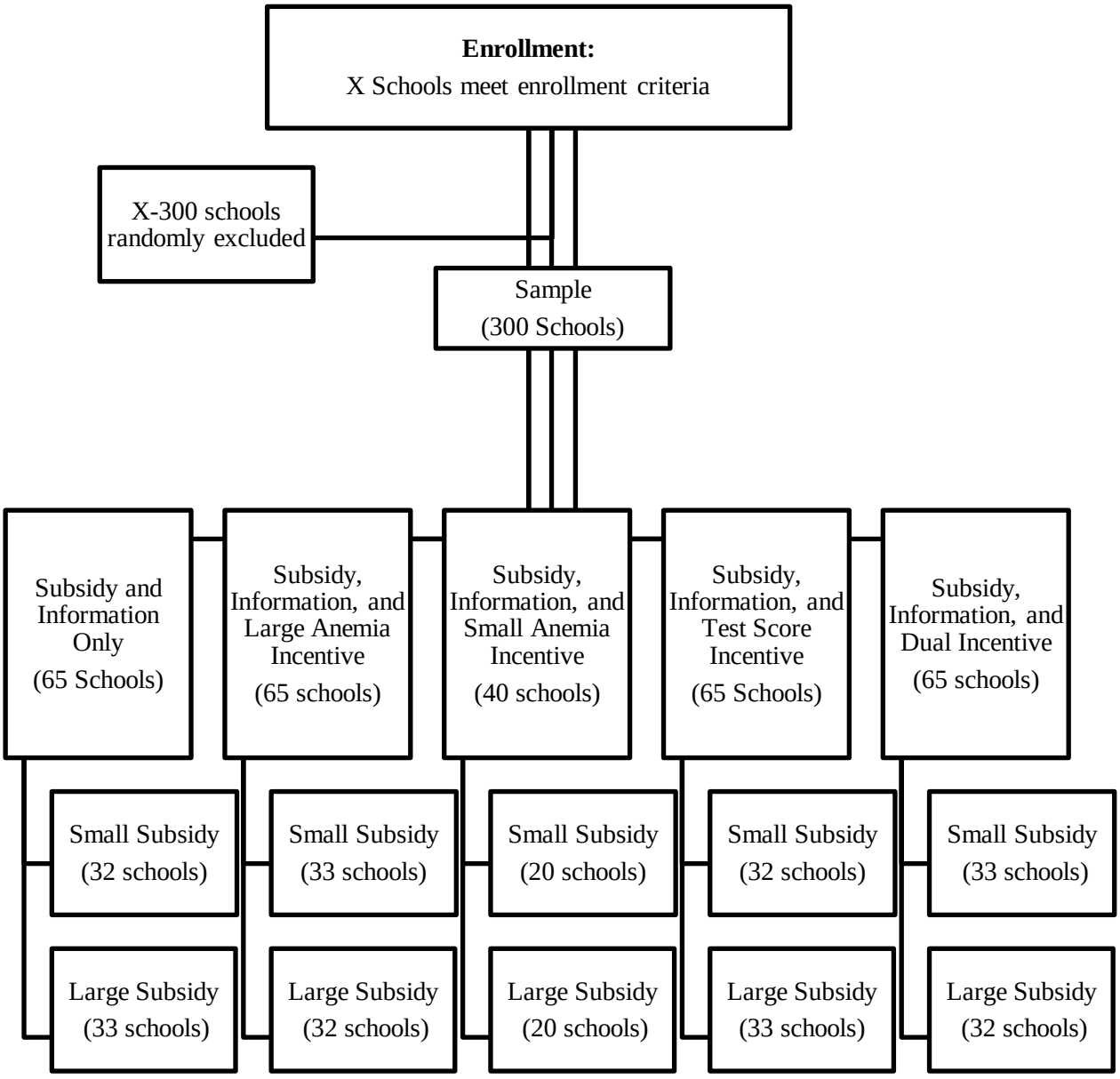
McKenzie, D. 2011. *Beyond Baseline and Follow-up: The Case for More T in Experiments*. World Bank.

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 : British Medical Journal* 345. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3406729/.

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

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

Appendix: Trial Profile



Appendix: Study Locations

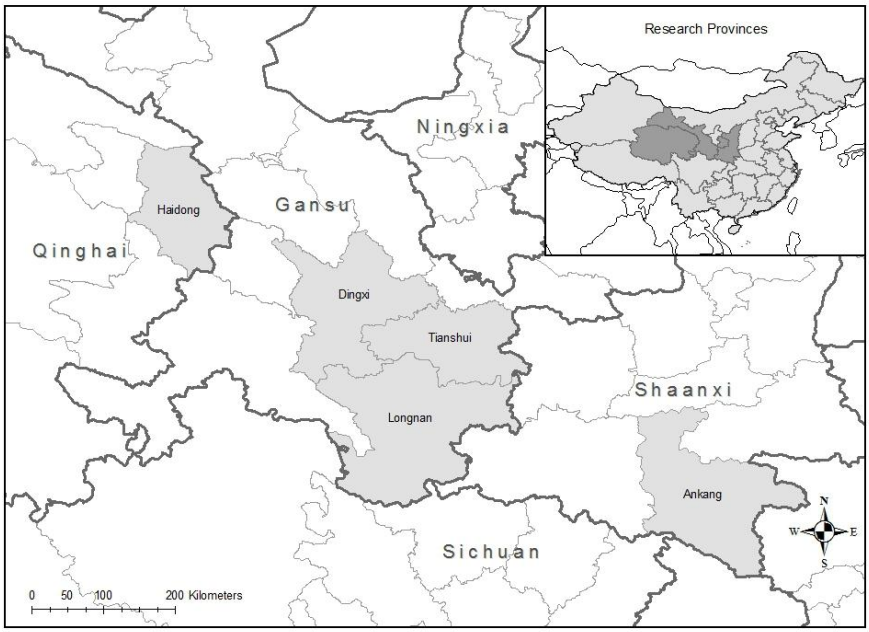


Figure 1: Study Regions

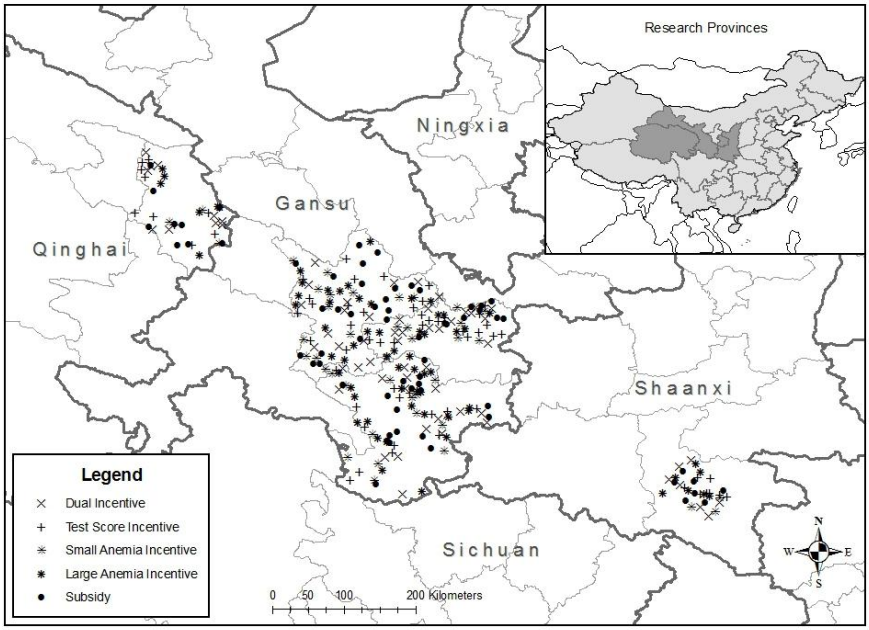


Figure 2: Study Sites