

Generasi Analysis Plan
Wave III

Benjamin Olken, Junko Onishi, and Susan Wong
January 22, 2010

This document describes the analysis plan for the *Generasi* Wave III survey. Note that this analysis document was written before looking at any of the Wave III data. All of the variables that were present in the Wave II survey retain the same variable definitions used in the Wave II survey, which were defined prior to looking at the Wave II data. More details of the Wave II analysis plan can be found in the Wave II analysis plan dated April 8, 2009.

Overall outline of paper / analysis:

0. Key metrics for Wave III analysis
 - 0.1. Analysis of key impact metrics specified ex-ante where we hypothesize *Generasi*'s effects will be most pronounced (section 1.2 below)
 - 0.2. Analysis of key final outcomes (section 1.3 below)
1. Impact: What is the overall effect of the *Generasi* program? What are the marginal effects of the incentives?
 - 1.1. What did program spend the money on?
 - 1.2. On targeted indicators
 - 1.3. On final outcomes
 - 1.4. On non-targeted indicators
2. Interactions: Where are the program's effects largest?
 - 2.1. Areas: Areas where services were low before, where access was difficult, where capacity to improve is low
 - 2.2. Changes over time.
 - 2.3. Individuals: For which types of individuals does *Generasi* have largest impact? People who didn't have much access before, poor, etc.
3. Mechanisms: Why do the program / incentives have an effect?
 - 3.1. Supply: Provider quantity
 - 3.2. Supply: Provider quality (health and education infrastructure quality)
 - 3.3. Supply: Provider effort
 - 3.4. Community effort at service provision and monitoring
 - 3.5. Price theory analysis: supply vs. demand shifts
4. Testing incentive theory: What are the potential costs and benefits of incentives?
 - 4.1. Learning / experimentation / efficiency of fund allocation
 - 4.2. Tighter targeting towards more 'marginal' individuals
 - 4.3. Price effects
 - 4.4. Corruption and sabotage
5. Organization of an academic paper on the incentives

Background

Improving the health and education of children is considered critical to economic development and forms an important component of the Millennium Development Goals. Faced with these challenges, many developing countries have sought to stimulate demand for maternal and child health services and education through conditional cash transfer programs. Mexico's Progresa program (Gertler 2004; Schultz 2004; Rawlings and Rubio 2005) for example, links cash payments to behaviors such as immunizations, growth monitoring, school enrollment, and school attendance. However, these types of demand-side interventions may be inappropriate in many developing world contexts, where beneficiaries do not have adequate access to health and education services (Schubert and Slater 2006, Lagarde, Haines, and Palmer 2007). In such environments, programs that address both the supply and demand side constraints may be more appropriate.

In 2007, the Government of Indonesia launched a large pilot of the Conditional Cash Transfer program applying two different approaches: conditional cash transfers to households and conditional cash transfers to communities. These two pilot projects are being implemented in six provinces, and are designed to achieve the same objectives and goals, in line with the Indonesian Government's priorities and the Millennium Development Goals: to reduce poverty; to reduce maternal mortality; to reduce child mortality, and to ensure universal coverage of basic education.

The **Household CCT** version, *Keluarga Harapan Project* (PKH) applies the traditional CCT design with quarterly cash transfers to poor individual households identified through statistical means. CCT recipient households receive regular cash transfers through the post office as long as they meet the requirements of using specified health and education services.

The **Community CCT**, known as *Generasi*, differs from the Household CCT in that block grants will be allocated to communities, rather than to individual targeted households. Under the program, over 1,600 villages received an annual block grant, which each village could allocate to any activity that supported one of 12 indicators of health and education service delivery (such as prenatal and postnatal care, childbirth assisted by trained personnel, immunizations, school enrollment, and school attendance). To give communities incentives to focus on the most effective policies, the government bases the size of the village's *Generasi* block grant for the subsequent year partly on the village's performance on each of the 12 targeted health and education indicators. The *Generasi* program thereby takes the idea of performance incentives from conditional cash transfer programs and applies it in a way that allows communities the flexibility to address supply constraints, demand constraints, or some combination. To the best of our knowledge, the *Generasi* program is the first health and education program worldwide that combines community block-grants with explicit performance bonuses for communities.

To allow for a rigorous, randomized evaluation of *Generasi*, the government of Indonesia incorporated random assignment into the selection of *Generasi* locations. Unlike evaluations of conditional cash transfer programs, which cannot separately identify the impact of the incentives from the impact of the additional cash provided (Gertler 2004), the *Generasi* evaluation was designed to separate out these two effects. Specifically, each *Generasi* location was further randomly allocated to one of two versions of the program: one "incentivized" version with the pay-for-performance component described above, and a second, otherwise identical "non-incentivized" version without the pay-for-performance incentives. This document describes the analysis plan for the first post-treatment wave of *Generasi*.

0. Key Metrics for Wave III analysis for Policy Purposes

- We will present to government / donors the impact on the 12 targeted indicators (Section 1.2), separately as well as combined.
 - We will present the Wave II and Wave III outcomes, as well as the average outcome. We will also present these separately for the incentivized and non-incentivized versions of the program.
 - For purposes of “whether the program should be continued” the government will primarily examine the Wave III outcomes, since these are the outcomes that best assess the current state of the project.
 - For the purposes of doing cost-benefit analyses, we will use the average Generasi effect (which is a weighted average of the Wave II and Wave III outcomes).
- We will present to government / donors the impact on the final outcomes indicators (Section 1.3), separately by wave as well as an average.

1. Impact.

Organization of this section:

- 1.1: What did program spend money on (direct benefits, old 2.3, and)
- 1.2: What impact did that have on service use (the indicators, old 1.1)
- 1.3: What impact did that have on ultimate outcomes we care about (health and education outcomes, old 1.2)
- 1.4: What impact did that have on other indicators that we weren't targeting (spillovers, old 1.3)

For each of the outcomes below, we will examine:

- Overall Generasi impact (randomized)
- Impact of incentives relative to non-incentives (randomized)

1.1. Direct benefits

- Outcome variables to examine
 - Survey-based measures of benefits
 - Health
 - PMT received at school (DLA25i)
 - PMT received at Posyandu (POS17)
 - PMT received intensively (POS18 anything once a week or more)
 - Subsidies received for health [amounts received transportation cost or service fees during pregnancy (CH42)
 - Subsidies received for delivery (CH50) for deliveries in the last 18 months
 - Education
 - Scholarships received for education (DLA25a,d,e)
 - Distribution of uniforms (DLAc,)
 - Other school supplies (DLA25b,f)
 - Transport (DLA25g,)
 - Other school (DLAh, DLA25v)
 - MIS-based measures of program expenses (note: this is only for comparing Versi A with Versi B)
 - Overall aggregate measures
 - All health expenditures
 - Health durables (e.g., infrastructure and furniture, health equipment – there should be less in Versi A since more premium on things with more rapid returns)
 - Health benefiting providers (e.g., expenditures on provider salaries, furniture / uniforms for posyandu workers – Versi A should have less of this)
 - Health expenditures detail
 - Subsidized care
 - Nutrition supplements
 - Drugs and health equipment
 - Wages and transportation of personnel
 - Infrastructure and furniture
 - Education expenditures detail
 - Scholarships
 - School uniforms
 - Training and outreach
 - Wages and transportation of personnel

- Infrastructure and furniture

1.2. Impact on targeted outcomes

- Outcome variables to examine
 - Health behaviors:
 - *Prenatal*: Number of pre-natal visits by all moms who gave birth in last 18 months
 - *Delivery*: Delivery by trained midwife/doctor, for all moms who gave birth in last 18 months.
 - *Postnatal*: Number of post-natal visits within 42 days after delivery by all moms who gave birth in last 18 months.
 - *Iron*: Number of iron tablet sachets during pregnancy for all moms who gave birth in last 18 months.
 - *Immunizations*: Percent of immunizations you should have had up to 11 months, for all kids 23 months old and below.
 - *Weight checks*: Number of weight checks in past 3 months, for all kids below age 3. We will use mom's recall of # posyandu visits in last 3 months (POS05), but 0 if child was not weighed at last visit.
 - *Vitamin A*: Number of Vitamin A supplements in past 18 months, for all kids above 6 months and below age 2.
 - Health:
 - *Weight*: % malnourished (< 2 sd) , all kids below age 3
 - Education:
 - *7-12 participation rate*: Enrollment dummy for age 7-12 in current school year (2009/2010 in Wave III, 2008/09 in Wave II vs. 2006/2007/2008 in Wave I). This comes from the 'are you in school' question on Form 1C (DLA09).
 - *13-15 participation rate*: Enrollment dummy for age 13-15 in current school year (2009/2010 in Wave III, 2008/09 in Wave II vs. 2006/2007/2008 in Wave I) (DLA09).
 - *7-12 attendance rate*: Percent of school days attended in last 2 weeks for age 7-12 from parents' report. This includes kids who are not enrolled.
 - *13-15 attendance rate*: Percent of school days attended in last 2 weeks for age 13-15 from parents' report. This includes kids who are not enrolled.
- *Additional Analysis To Be Reported in Tables*
 - Will also include the following additional education variables, although we will not include them in the calculation of average standardized effects
 - Age 13–15 conditional attendance
 - Age 13–15 enrolled in SMP
 - Age 7–12 enrolled in SD
 - Age 13–15 enrolled other than SMP
 - Kecamatan SD gross enrollment
 - Kecamatan SMP gross enrollment
- Additional analysis of interest : weighted average of above outcomes using Generasi program weights.
 - First, we fix the average 'jumlah sasaran' for each indicator for a particular village. We need to weight by average number of sasaran because regressions will project for each variable the increase of Generasi per sasaran, so to get the number of points you need to multiply by the average number of sasaran
 - Then we use the Generasi Bobots to aggregate the 'predicted increase in score' that a village would get under Generasi.

- Specifically, we run all 12 of these regressions simultaneously to get the “Generasi” regression coefficient for each of these 12 indicators. You would then construct the “Total Generasi Effect”, and test the null hypothesis that $\beta_1 w_1 \theta_1 + \beta_2 w_2 \theta_2 + \dots + \beta_{12} w_{12} \theta_{12} = 0$ where β_i is the regression estimate for indicator i and $w_i \theta_i$ is the bobot * average # of sasaran (i.e., the weight in the table above.)
- See table below for an example of the weights:

Generasi Indicator From Form 17	Generasi Scoring Weight Per Indicator	Survey Indicator	Adjusted weight for regression variable to equal annual Generasi bobots	Average # Sasaran in a village in 12 months. Note these are from the MIS, assuming an average person had 1000 total sasaran; they can be re-run with the final MIS numbers.	Weight for regression coefficient
1. Prenatal care visits	12	Prenatal care visits during pregnancy	12	75	900
2. Iron pill receipt (30 pill supply)	7	Iron pill receipt (30 pill supply) during pregnancy	7	75	525
3. Delivery assisted by a trained professional	100	Delivery assisted by a trained professional	100	75	7500
4. Postnatal care visit	25	Postnatal care visit during pregnancy	25	75	1875
5. Immunization coupon	4	Immunizations	4	75	300
6. Monthly weighing with weight increase	4	Underweight	48	295	14160
7. Monthly weighing visit (under 3)	2	Number of weighing visits in last 2 months	12	285	3420
8. Vitamin A pill	10	Vitamin A pill receipt in last year	10	75	750
9. Enrollment SD	25	Enrollment SD	25	444	11100
10. Month with > 85% attendance SD	2	85% attendance SD in last 2 weeks	24	444	10656
11. Enrollment SMP	50	Enrollment SMP	50	184	9200
12. Month with > 85% attendance SD	5	85% attendance SD in last 2 weeks	60	184	11040

- Additional notes

- Attendance:
 - The main analysis uses parents’ report of attendance, rather than the school based check.
 - We also observe attendance directly through random spot-checks at school. Based on our analysis of the baseline data, we concluded that this metric has lower power (since we only observe one class per school), but we will use this variable as an alternative check.
- Immunizations:
 - A potential issue is that immunization record cards are more likely in Generasi locations. This could lead to differential accuracy in reporting of immunization status in Generasi vs. control areas.
 - We will therefore:
 - Check this by first examining whether the probability mom has record card differs in Generasi vs. control

- If so, look at whether probability of correctly recalling the BCG vaccine (i.e., compare mom answer to scar) is greater in Generasi locations vs. control to gauge whether differential recall is a problem
- If both the probability of having a card differs in Generasi vs. control and there is differential recall of BCG scar in Generasi vs. control, we will note that it should be interpreted with caution for this reason, and a compute a version of the 'average standardized effect' excluding immunization as well as including it.

1.3. Impact on final outcomes.

- Key outcome variables to examine
 - Health:
 - Morbidity
 - Malnourished (Z score less than 2) (for all under 3s)
 - Severe malnourished (Z score less than 3) (for all under 3s)
 - Wasting (weight for height Z score less than 2) (for all under 3s)
 - Severe wasting (weight for height Z score less than 3) (for all under 3s)
 - Stunting (height for age Z score less than 2) (for all under 3s)
 - Severe stunting (height for age Z score less than 3) (for all under 3s)
 - Acute illness in past month (for all under 3s). (Dummy variable for having had either Diarrhea or ARI in past month).
 - Mortality
 - Neonatal mortality (0-28 days, all births in last 18 months)
 - Infant mortality (0-11 months, all births in last 24 months)
 - Education:
 - Home-based test scores (age-adjusted Z-scores, where we compute the distribution in the control areas)
 - Math
 - Bahasa
 - Total (not included in avg std effects to avoid double counting)
- *Additional Analysis (not included in average standardized effects)*
 - Nutrition variables (all 6 above) broken down by:
 - Infant (0-11 months)
 - 1-3 year old
 - School-based test scores. We will not group with other education scores, given that if we enroll marginal kids, this could cause average test scores to fall.
 - Mean UAS score
 - Mean UN score
 - Consumption expenditure. We do not expect Generasi to affect consumption, but HH CCT might affect consumption, so we will need to look at it just to be sure.
 - Per capita consumption
 - Per capita consumption broken down by initial consumption quintile (panel HH only)
 - Consumption quintiles will be based on our baseline survey (weighted.)
 - Motor development – age in months at: sitting alone (BM02a); crawling (BM02b); standing with support (BM02c); walking with support (BM02d); standing alone (BM02e); and walking alone (BM02f)
-
- Note that avg std effects is the avg of the morbidity and mortality effect, counting each of those equally
- Note:
 - Height variables: We will use the baseline value of wasting and stunting, but code as missing in the Wave I data (i.e., treat as cross-section) any households that are flagged as likely measurement error by the Z score program.
 - Note: need to check for differential response rate to home-based tests in treatment and control
 - Keep “mortality 0-12 months among births in past 12-24 months” as an “additional indicator by interest”

1.4. Report supplemental additional analysis where all the nutrition variable are broken down by ages

Impact on non-targeted indicators

- Outcome variables to examine
 - Health service provision and use:
 - Quality of prenatal care services - completeness of content for most recent pregnancy (if most recent pregnancy within 18 months) (what share of following services: measure weight (CH15a), measure height (CH15b), blood pressure (CH15c), sampled blood (15d), measured waist circumference (CH15e), check position of the fetus (CH15f) internal inspection (CH15g), and measured hip circumference (CH15h), discussion about potential complication (CH16), TT shot (CH18))
 - Facility-based deliveries vs home deliveries (percent of deliveries in the last 18 months that were done in a facility – CH24 anything other than ‘rumah dukun bayi and rumah sendiri/rumah keluarga’ counts as facility)
 - Use of Family Planning (percentage of women age 15-45 who answer ‘yes’ to KB02, use a modern method (types 1-9 on KB03), and for those using types 1-5 have received contraceptives sometime within the past 6 months)
 - Use of health services for curative care (percent of those who used modern health services (not self-medicated or traditional provider) among household members who were sick in the past one month (sick - MA04 and use - RJ04 and RJ06)
 - Quality of Posyandu: Any communication with the mother at last visit about health status of kid (POS12 and POS13)?
 - Parental behaviors and knowledge:
 - Initiation of breastfeeding (did you start within 1 hr of delivery?)
 - Exclusive breastfeeding (up to 3 months, na05 and na06 \geq 4 months or more), all kids born in last 18 months
 - Mother’s knowledge on exclusive breastfeeding and treatment of diarrhea (percent of knowledge questions answered correctly, for moms with children under 3)
 - Child labor. (lower number is better)
 - Number of hours child age 7-15 worked in wage work in last week.
 - Number of hours child age 7-15 worked in households work in last week.
 - Education (other enrollment margins)
 - Age 16-18 school participation rate (dummy for whether enrolled in school for 16 – 18 year olds, AR10 and AR11 in buku 1A).
 - Drop out rates (dummy for whether child dropped out from school in the last 2 school years DLA01, DLA04 and DLA05 in buku 1C or DS17 and DS17a in buku 5).
 - SD to SMP transition (you were in SD two years ago, you should have been in SMP now, did you make it).
 - Number of hours attended school in last week – DLA 19/20
 - Number of hours attended school in last week, conditional on being enrolled – DLA 19/20
 - Numbers attending Paket A (non-formal education for SD) – age 6 – 22 (Book 1A AR53-54)
 - Numbers attending Paket B (non-formal education for SMP) – age 6 – 22 (Book 1A AR53-54)
 - Numbers attending Paket C (non-formal education for SMA) – age 6 – 22 (Book 1A AR53-54)
- Note that I’d like average standardized effects here to be grouped as above:
 - Average health
 - Health service provision
 - Parental health behavior
 - Average education
 - Child labor
 - Other enrollment margins

- Additional variables to examine (not for standardized effects):
 - Fertility
 - Fertility rate (defined as percent of women age 15-45 who gave birth in last 12 months). This we do not expect to be an effect, so it shouldn't be included in the average indicators,
 - Migration
 - Average number of people migrated out from village within the past 12 months (village level aggregate, Book 2 ID11)
 - Number of individuals within households migrated to outside of kecamatan within the last 12 months (Book 1A, AR17)
 - Excess health (not main indicator but could look at)
 - Excess prenatal visits (% who received >4 visits)
 - Excess Vitamin A (% who received >100%)
 - Transport to school (sign is unclear: program could fund closer schools, or program could subsidize transport to further better schools)
 - Distance to SMP attended in kilometers (buku 1C DLA13)
 - Time spent one way to SMP (buku 1C DLA14)
 - Transportation cost one way to SMP (buku 1C DLA16)
 - Women's decision making power subjective question (not a main indicator):
 - Dummy for whether women has a role in each of the SP01 answers (education, health, discipline, fertility)
 - dummy for woman saying 'no' for permission in SP02 answers (buying food, clothing, medicine, personal goods)
 - Education
 - Transition alternate definition: In the alternate definition, kids who report having the last grade as klas 6 SD, not graduated, are counted in the denominator and considered as non-transitioning for the numerator. In the first version, these klas 6 non-graduating kids are not included.

2. Where are program effects largest?

This analysis examines heterogeneity in the program's impact. We examine changes to the program over time, differences by province, differences by income level, and differences based on pre-existing conditions. We focus on:

- What the program spent the money on, shown above in Section 1.1.
- The 12 targeted indicators listed above in Section 1.2.
- The final outcomes listed above in Section 1.3

Note: since these are interactions, and not explicitly part of the randomization design, this is the part it is most important to specify completely ex-ante.

2.1. In which areas is the program most effective?

- Key heterogeneity to look at:
 - Pre-period level:
 - For each indicator, look at interaction with pre-period levels to see if there are bigger impacts in places with lower levels of performance at the kecamatan level. We will report program effects at the 10th percentile of baseline level performance.
- Other analyses that are of interest to the government and will be explored:
 - Java/NTT/Sulawesi:
 - For each of sets of indicators above, look at interaction with Java/NTT/Sulawesi. We will report the program effects for Java, NTT, and Sulawesi, and tests for whether NTT = Java and whether Sulawesi = Java.
 - Pre-period general kecamatan poverty:
 - For all indicators, interact with log of average per-capita consumption of the kecamatan.
 - Pre-period village access variables:
 - For all health indicators, look at interaction with whether the village had a bidan or Puskesmas located in the village in the pre period
 - For SMP indicators, look at interaction with whether the village had an SMP in village in pre-period
 - (will not look at SD indicators for access, since all villages have SD)
- Other analyses of more academic interest:
 - Pre-period social capital measures:
 - Within kecamatan, are there bigger impacts in places with more social capital? Measure social capital by the average number of groups of all HH in the village except you. Include kecamatan FE so this is a within kecamatan analysis. This is of particular interest for the incentives analysis.
 - Education levels.
 - Do kecamatans where average education levels are higher respond more to the incentives?
 - Could be interesting to explore this both for the village head's education level in the pre period (does this help him get more points within the kecamatan) as well as education levels more broadly.

2.2. *How did the program change over time?*

- To examine changes in the program over time more precisely, we will limit the analysis to the 129 kecamatans that were Generasi in Wave II and Wave III) and the kecamatans that were control in both year (i.e. we drop the 49 kecamatans where the program started in Wave II).
- We will then repeat analysis 1.1, 1.2, and 1.3 on this sample, testing explicitly whether Generasi's impact changed between waves.

2.3 For which types of individuals does *Generasi* have the largest impact?

- Analyses to do
 - Pre-period poverty using panel household:
 - For each of the indicators above, conduct analysis separately for bottom 2 quintiles vs. top 3 quintiles. Consumption quintiles will be based on our baseline survey (weighted.)
- Interesting but not main analysis:
 - To examine elite capture: for each of the indicators in 1.1 above, as well as each of the 12 main indicators, look at interactions to see if there are bigger impacts for those individuals who reported to “kenal dekat” one of the aparat desa. Include village fixed effects.
 - Pre-period poverty interaction with village fixed effects. Note that we are not going to have power to do this in most cases except for education indicators. But very interesting for within-village poverty targeting analysis if we can do it. Also run a specification with consumption squared to see if incentives moves you towards the middle of the income distribution (i.e., where people who are most ‘marginal’ may be)
 - Interact using GPS distance from the kepala desa’s office to measure remoteness / isolation of household. Include kecamatan fixed effects.

3. Why do the program and the incentives have an effect? Teasing out mechanisms.

The analysis in this section explores a varied of channels through which the Generasi program as a whole, and the incentives in particular, could have impacts.

3.1. Supply: Provider quantity

- Outcomes to examine:
 - Midwife
 - Presence of midwife having regular practice in village
 - Posyandu
 - Number of posyandu in village (from village head questionnaire)
 - Education
 - Presence of SD in village (including satu atap, terbuka, klas jauh from Buku 2)
 - Presence of SMP in village (including satu atap, terbuka, klas jauh from Buku 2)
 - Number of teachers in SD (include all teachers including part time / honor teachers)
 - Number of teachers in SMP (include all teachers including part time / honor teachers)

- Additional things to look at (we don't expect effects, but could be interesting just to know how these variables change)
 - Puskesmas.
 - Number of full-time health personnel (excluding admin and support staff)
 - Number of all full-time and part-time health personnel (excluding admin and support staff)
 - Number of midwives
 - Total midwife to population ratio, where we hold the population variable constant using the Wave I Puskesmas population number
 - Education
 - Number of teachers in SD (include only full time teachers)
 - Number of teachers in SMP (include only full time teachers)

3.2. Supply: provider inputs

- Outcomes to examine:
 - Midwife
 - Infrastructure
 - Access to water at location of practice (has access to closed water source (PAM, pump, well) within 10 meters of building IR18, IR19 and IR20 buku 4)
 - Has electricity at location of practice
 - Stock of basic essential drugs (Percent of essential drugs in stock at time of interview: Amoksisilin 250 mg (OV3d) Amoksisilin 500m (OV3d) Amoksisilin sirup (OV3f), Antalgin 500mg (OV3i), Parasetamol Sirup (OV3k), Parasetamol 500mg (OV3m))
 - Percent of tools they have: (index: Blood pressure measure, Forcep, Vaginal speculum, Tenakulum, Uterus sound, Gynecologist table, straight or curved clamps, weighing kit, and vaccine carrier buku 4)
 - Percent with stock of oxytocin (OVTYPE a1)
 - Quality of services: proportion of last three deliveries using partograph (YK08a)
 - Percent “always do” in antenatal care service items in public practice (PK01a)
 - Percent “always do” in antenatal care service items in private practice (PK01b)
 - School
 - Infrastructure – number of classrooms (DS08a)
 - Condition of infrastructure – index condition of chairs and desks for students, condition of the floor, condition of the walls, and condition of the roof. We include all the condition variables that are directly observed.
 - Has latrine for students (OL24, buku 5)
 - Latrine for students have enough water (OL26, buku 5)
 - Compute average standardized effects for:
 - Health
 - Education
 - SD
 - SMP

Other things to examine:

- Puskesmas:
 - Stock out of any vaccine within last two months (BCG, Polio, measles, and DPT&HepB or DPT HepB Combo buku 3)
 - Note that this is particularly more prevalent in NTT. So we will want to do provincial analysis split on this indicator.
 - Note also that our power calculations suggest that it is going to be very hard to detect effects here – given that the mean is only 10% of puskesmas are stocked out of any vaccine overall and only 20% of puskesmas are stocked out of any vaccine off Java.

3.3. Supply: provider effort

- Outcomes to examine
 - Health:
 - Midwives
 - Midwife hours in last 3 days:
 - Providing neighborhood outreach
 - Providing public services in office
 - Providing private services
 - Total hours worked
 - Number of posyandu sessions a midwife attended in the last one month
 - Number of hours spent by midwives per posyandu
 - Education:
 - Teacher absence. Percent of teachers who are present at time of interviews.
 - Teacher observation. Percent of time teacher is teaching (OL27) at time of interview
- Additional notes – could be worth investigating, but we don't expect to see much on these variables.
 - Puskesmas:
 - Minutes wait at recent health visits
 - Absence of providers (percent of providers on list who are there at the time). Note that if there is lower absenteeism, it's likely a good thing, if higher absenteeism, need to look at whether spending more time in the field.

3.4. Community effort at service provision and monitoring

- Outcomes to examine:
 - Conceptually we can divide these into three types of ‘community effort’
 - Community effort at direct service provision.
 - Number of active posyandu in village (DN07, sum of all dusuns)
 - Number of posyandu meetings in past year at selected posyandu (Buku 6, IDP03)
 - Number of kaders posyandu at selected posyandu (Buku 6, IDP04)
 - Community effort at outreach (going around and pressuring people to make sure that they complete services, socializing Generasi, socializing the importance of health and education, planning activities.)
 - Number of ‘sweepings’ in last year (Buku 6, IDP06)
 - Number of school committee meetings with parents during past school year (Buku 5, MS16c)
 - Community effort at monitoring (community making sure service providers are doing their job)
 - Number of school committee members (Buku 5, MS14)
 - Number of school committee meetings with teachers during past school year (Buku 5, MS16b).
 - Participation in health / education programs
 - For Incentive/Nonincentive in Wave II and for all analysis in Wave III: participation in meetings about health education (will be non-Generasi specific in Wave III)
 - Proportion of kids under 3 who own buku kupon (PG15, PG 16, PG17, buku 1A)
 - Proportion of kids under 3 with buku kupon with evidence of use (coupons stamped and/or collected buku 1B and buku 1D)
 - Proportion of kids under 3 who have Buku KIA/ KMS
 - Proportion of households that think health services in general has improved over the last two years (PM29 – 1)
 - Proportion of households that think education in general has improved over the last two years (PM29 – 1)
 - Spillovers to other types of community activities
 - Participation in gotong royong (number of person-hours from the household)
 - Participation in women’s groups (number of meetings)
 - Participation of women respondent in activities of type F (number of meetings)
 - Overall participation in social groups (number of meetings)
 - Participation in general election 2009
 - Average standardized effects to be reported:
 - Total
 - Community effort
 - Direct provision
 - Outreach
 - Monitoring
 - Participation in health/education programs
 - Participation in other types of community activities
- Additional notes:
 - We will work to improve these indicators in Wave III.

3.5. Price theory analysis: supply vs. demand shifts, who gets rents

We can think of *Generasi* as affecting two different margins – the supply of services and the demand for services. Even though we've seen (hopefully) changes in quantities or service provider behavior, that doesn't mean that it was supply necessarily – that could also be due to a change in demand. By looking at prices and quantities jointly we can say something about shifts in supply and demand curves. For example, if price increases and quantity increases, we know that demand shifted out since a supply shift cannot cause a simultaneous price and quantity increase holding the demand curve constant.

For this analysis, then, we want to examine prices and quantities using comparable metrics.

- Variables to examine
 - Midwife services
 - Normal childbirth at private practice
 - Fee charged by midwife (buku 4)
 - Quantity done by midwife in last month (buku 4)
 - Normal childbirth at government practice
 - Fee charged by midwife (buku 4)
 - Quantity done by midwife in last month (buku 4)
 - Normal childbirth by midwife (combined of private practice and public practice)
 - Fee charged by midwife (buku 4)
 - Quantity done by midwife in last month (buku 4)
 - Fee paid by mother (non-Askeskin holders)
 - Antenatal care services at private practice
 - Fee charged by midwife (buku 4)
 - Quantity provided by midwife in last month (buku 4)
 - Antenatal care services at government practice
 - Fee charged by midwife (buku 4)
 - Quantity provided by midwife in last month (buku 4)
 - Antenatal care services at (combined of private and government practice)
 - Fee charged by midwife (buku 4)
 - Quantity provided by midwife in last month (buku 4)
 - Fee paid to midwife (buku 1B) (non-Askeskin holders)
 - Family planning (3 month shot) at private practice
 - Fee charged by midwife (buku 4)
 - Quantity done by midwife in last month (buku 4)
 - Family planning (3 month shot) at government practice
 - Fee charged by midwife (buku 4)
 - Quantity done by midwife in last month (buku 4)
 - Family planning (3 month shot) at (combined private and government practice) – confirmed most common form of FP
 - Fee charged by midwife (buku 4)
 - Quantity done by midwife in last month (buku 4)
 - Fee paid to midwife (buku 1B) (non-Askeskin holders)
 - Puskesmas services
 - Normal childbirth at Puskesmas assisted by Midwife
 - Fee charged (buku 3)
 - Quantity done in last month (buku 3, note that qty doesn't distinguish doctor and midwife)
 - Fee paid (household survey) (non-Askeskin holders)

- Posyandu services
 - Fee for posyandu visit (buku 6)
 - Quantity of kids seen at posyandu in last month (buku 6)
- SD
 - Annual cost of school for TA 07/08 (from buku 5)
 - Number of students enrolled in TA 07/08 (from buku 5)
 - Number of students enrolled in TA 08/09 (from buku 5)
 - Cost of school from parents for previous semester
- SMP
 - Annual cost of school for (TA 07/08 Wave II; TA 05/06 Wave I) (from buku 5)
 - Number of students enrolled in (TA 07/08 Wave II; TA 05/06 Wave I) (from buku 5)
 - Number of students enrolled in (TA 08/09 Wave II; TA 05/06 Wave I) (from buku 5)
 - Cost of school from parents for previous semester
- Average standardized effects reported:
 - Fees
 - Health
 - Midwife
 - Village health post
 - Puskesmas
 - Education
 - SD
 - SMP
 - Quantities
 - Health
 - Midwife
 - Village health post
 - Puskesmas
 - Education
 - SD
 - SMP

4. What are the potential costs and benefits of incentives?

This section tests for some of the positive and negative impacts of the incentives in the *Generasi* program. The analysis is therefore focused on comparing Generasi Versi A (incentives) with Generasi Versi B (non-incentives).

4.1. Learning, experimentation, and efficiency of *Generasi* fund allocation

The general theory is that incentives encourage the community to allocate their funds more efficiently. Moreover, incentivized communities learn more about efficient allocations over time.

- Analysis to do (this is covered in 1.1 above):
 - Overall changes in categories of expenditure (MIS data)
 - Overall aggregate measures
 - All health expenditures
 - Health durables (e.g., infrastructure and furniture, health equipment – there should be less in Versi A since more premium on things with more rapid returns)
 - Health benefiting providers (e.g., expenditures on provider salaries, furniture / uniforms for posyandu workers – Versi A should have less of this)
 - Health expenditures detail
 - Subsidized care
 - Nutrition supplements
 - Drugs and health equipment
 - Wages and transportation of personnel
 - Infrastructure and furniture
 - Education expenditures detail
 - Scholarships
 - School uniforms
 - Training and outreach
 - Wages and transportation of personnel
 - Infrastructure and furniture

Note that this is the one set of results where we have already looked at the data from Year 1:

<i>Share of block grant on:</i>	Mean incentives	Mean no incentives	No fixed effects		District fixed effects	
			Treat. effect	p-value	Treat. effect	p-value
All health expenditures	0.458 (0.012)	0.420 (0.019)	0.038 (0.022)	0.095*	0.035 (0.015)	0.026**
Health durables	0.084 (0.012)	0.074 (0.014)	0.010 (0.019)	0.588	0.017 (0.013)	0.188
Health benefiting providers	0.095 (0.014)	0.088 (0.014)	0.007 (0.020)	0.731	0.016 (0.014)	0.251
<i>Health expenditures</i>						
Subsidized care	0.106 (0.010)	0.124 (0.012)	-0.017 (0.016)	0.271	-0.007 (0.010)	0.477
Nutrition supplements	0.219 (0.013)	0.182 (0.013)	0.037 (0.019)	0.049**	0.019 (0.014)	0.177
Drugs and health equipment	0.016	0.009	0.007	0.067*	0.005	0.100*
Wages and transportation of Personnel	0.036 (0.005)	0.032 (0.005)	0.004 (0.007)	0.554	0.004 (0.006)	0.505
Infrastructure and furniture	0.081	0.074	0.007	0.725	0.014	0.286

<i>Education expenditures</i>	0.156	0.153	0.003	0.920	0.008	0.723
Scholarships	(0.017)	(0.022)	(0.028)		(0.022)	
School uniforms	0.247	0.307	-0.060	0.016**	-0.058	0.011**
	(0.016)	(0.019)	(0.025)		(0.023)	
Training and outreach	0.014	0.016	-0.002	0.604	-0.001	0.568
Wages and transportation of Personnel	0.028	0.029	-0.001	0.902	0.001	0.853
	(0.005)	(0.005)	(0.007)		(0.006)	
Infrastructure and furniture	0.097	0.075	0.023	0.111	0.016	0.136
	(0.011)	(0.009)	(0.014)		(0.011)	

- Additional things to investigate:
 - Increasing “predicted points” of expenditure.
 - Increasing average predicted points. Prediction is that Versi A kecamatans should be focused on those expenditures that in general produce more points.
 - Run regression of “points” on dollars in Versi B kecamatans, using the MIS data, with kec FE and jumlah sasaran as RHS variables. Note that we include all points *including the minimums*, not the points after having subtracted the minimums. This coefficient is the OLS relationship between a dollar of spending in a particular category on total points.
 - This yields a ‘weight’ for each type of expenditure – how many ‘predicted points’ it generates per dollar. We can then use these weights to assign a ‘total predicted points’ to each village based on their spending profile.
 - We then run regressions of
 - Efficiency: Predicted points on Versi A (i.e., do Versi A choose expenditures with higher predicted points) in year 1
 - Efficiency: Predicted points on Versi A (i.e., do Versi A choose expenditures with higher predicted points) in year 2
 - Note: do we want to do this province specific or even kabupaten specific? That would allow the ‘efficient’ expenditures to vary by location. Or something else?
 - Experimentation:
 - Do budgets change more from year 1 to year 2 in Versi A vs. Versi B kecamatans? The prediction is that they change more in Versi A as villages experiment more in order to find the optimal allocation of funds.
 - Compute sum of squared differences in expenditure shares. I.e., for each of the , i.e., for each of the 10 detail expenditure categories, calculate $CHANGE_I = ABS(YR2SHARE - YR1SHARE)$. Then calculate $TOTALCHANGE = SUM(CHANGE_I)/2$.
 - This is an index from 0-1 describing what fraction of budget allocations were reallocated from year 1 to year 2.
 - Regress whether TOTALCHANGE is higher in Versi A kecamatans.
 - Learning about efficiency:
 - Change in predicted points (calculated using “overall efficiency” as above) from year 1 to year 2 on Versi A (i.e., do Versi A change their allocations more to increase predicted points more (i.e., do they learn more)
 - Maximizing rewards given nonlinearities in the point system
 - Given the minimum thresholds you should *concentrate* your spending to get over minimum thresholds.
 - Compute the Herfindahl of spending on the 10 categories, and see whether spending is more concentrated in Versi A than Versi B.
 - A slightly more sophisticated version is that you should focus on areas where marginal points are highest. The points system, with minimum thresholds, implies that you should

focus on a) Those items where you have enough baseline performance that most of the expenditure will be marginal (i.e., generate points) and b) those items where your performance is not already so high that there is little room for performance. This predicts that the incentives lead to an inverted U-share relationship between baseline performance on an indicator and spending on items that improve that indicator. To test this we will:

- In Versi B kecamatans, for each of the 12 indicators defined in impact (1) above, regress the change in that indicator on that village's spending in each of the 10 categories listed above. This yields a 12*10 matrix θ , where θ_{ij} tells you how much a marginal dollar on input j impacts outcome i .
- The prediction is that spending to impact outcome i is an inverted U shape with respect to the baseline level on outcome i in Versi A, but not in Versi B.
- Therefore we take the spending vector S and calculate θS , which is a matrix of the predicted impact of that spending on all 12 indicators j . The analysis of overall efficiency is that the weighted sum of θS is higher in Versi A than Versi B. The analysis of specific efficiency predicts that if you regress:
 - $\theta S_j = \text{baseline} + \text{baseline}^2 + \text{baseline} * \text{VERSI_A} + \text{baseline}^2 * \text{VERSI_A} + \text{eps}$
you will get a negative coefficient on $\text{baseline}^2 * \text{VERSI_A}$.
 - This seems hard to really get – perhaps there is a better way of doing this?
- Learning about reward system:
 - Change in concentration of spending (herfindahl) from year 1 to year 2. Does concentration increase more in Versi A locations than in Versi B locations?

4.2. Tighter targeting towards more 'marginal' individuals

- Outcomes to look at
 - Direct benefits of Generasi funds – same indicators as Section 2.3 above.
- Analysis:
 - Using the control kecamatans, we'll run models to change in indicators conditional on per-capita consumption interacted with other household and beneficiary characteristics (age, lagged values, gender, etc). So we can compute the marginal effect of income for a given individual as a function of their characteristics.
 - With incentives, we predict largest effects for those who are closest to the threshold – i.e., for whom $d\text{Indicator}/d\text{Income}$ is greatest.
 - We then construct $d\text{Indicator}/d\text{Income}$ for panel households.
 - We'll then run a non-parametric regression of the change in takeup on the $d\text{Probability}/d\text{Income}$ score, interacted with the treatments, to see whether the treatments are more likely to change the outcomes for those closer to the threshold and, more generally, to examine the differential incidence of the program.

4.3. Price effects

- Price effects
 - Prediction: if the Rupiah value of a point is larger you work harder.
 - Methodology:
 - Calculate the predicted Rupiah value of a point.
 - Use baseline (year 0) number of sasaran and access variables and apply our susenas regressions and BLM formula to calculate predicted rupiah value of a point in the kecamatan. Make sure baseline number of sasaran values are not different incentive vs. non-incentive areas (if it is, instrument with the dusun sampling form.)
 - We then run a regression of achievement of:
 - 12 main indicators
 - Community effort variables
 - on predicted Rupiah value of a point interacted with incentive treatment, controlling for predicted Rupiah value and main effect of incentives.
 - Prediction: if you are more likely to be ‘in the money’ you work harder
 - Methodology:
 - Using the baseline # of sasaran for each category to predict likelihood of village being ‘in the money’ on a given indicator in year 1, using data from versi B locations. Make sure it’s not different incentive vs. non-incentive areas (if it is, instrument with dusun sampling form.)
 - For each of 12 indicators, interact ‘predicted in the money’ with incentives to see if those who are more likely to be in the money do better
 - Do the same for education and health on average (% of ed. indicators in the money, % of health indicators in the money)
 - Run the same regression on:
 - Allocation of funds to education vs. health (i.e., if you are more likely to be ‘in the money’ on education are you more likely to spend money there in treat relative to control)
 - Community effort on health (posyandu) vs education (school committees)

4.4. Corruption and Sabotage

- Outcomes to examine:
 - Are scores inflated overall?
 - For each kecamatan, use MIS data to compute % of target group that is achieving each of the 12 target indicators.
 - Then, for each kecamatan, compute the difference between the percent of HH achieving target indicator from MIS and the percent achieving according to survey
 - Regress difference on Versi A dummy
 - Are teachers inflating attendance?
 - Regress difference between recorded attendance on August 4 (random date for back check of attendance) and observed attendance on date of survey on Versi A dummy and also program as a whole.
 - Are number of sasaran inflated?
 - In Versi B the incentive is to only inflate sasaran; in versi A you have more an incentive to not inflate sasaran.
 - Key indicators:
 - Number of 'sasaran' in the program.
 - Change in sasaran from year 1 to year 2 (from MIS data)
 - Sabotage: do you exclude neighboring villages from service?
 - Midwives – # of posyandus you do outside your main place of practice
 - Household survey: # of people going to school outside their village and seeing a midwife located outside their village – does this go down in Versi A relative to Versi B
- Predictions:
 - Versi A will inflate scores, whereas versi B inflates # sasaran
 - There will be less inflation (of either type) in kecamatans with fewer villages, since there is more of an incentive to monitor each other
 - There is more inflation when the expected Rupiah value of an additional sasaran is higher

5. 5. Academic paper on the effect incentives.

Our analysis document of June 11, 2008 had the following structure, which still seems like a good set of guidelines for writing the academic paper on the impact of incentives. I have included the structure below, as well as a mapping to this document.



080611 MDG
Theoretical Proposal c

1. Does making aid conditional improve short-run performance on targeted indicators?
This maps to Section 1.2 of this document.
2. Does making transfers conditional transfers increase or decrease prices? Related to this, does making aid conditional succeed by increasing demand or increasing supply, or both?
This maps to Section 3.5 of this document.
3. How does making aid conditional change the within-village incidence of benefits?
This maps to Section 2.3 and Section 4.2 of this document.
4. Are non-incentivized intermediates complements or substitutes?
This maps to Section 1.4 of this document.
5. How does the program change the time horizon of village investments? How does the program change the types of village investments? Do they move towards investments that are likely to increase points?
This maps to Section 4.1 of this document.
6. How do incentives change the work behavior of health and education service providers?
This maps to Section 3.3 of this document.
7. Do incentives change the number and composition of people involved in village activities?
This maps to Section 3.4 of this document.
8. Do incentives exacerbate the persistence of temporary shocks?
9. What happens to final outcomes?
This maps to Section 1.3 of this document.

Important aspects not included in the above:

- Do the incentives lead to increased corruption and sabotage?
This maps to Section 4.4 of this document.
- Do villages respond to price effects embedded in the incentives?
This maps to Section 4.3 of this document.
- Do the incentives lead to increased learning on the part of villages?
This maps to Section 4.1 of this document.

Regression specifications.

- For all analysis, we run two regressions:
 - GENERASI vs. CONTROL. This captures the overall program effect of interest for most regressions, and lumps the incentivized and non-incentivized versions of the program together for maximum statistical power. For the overall policy evaluation, this is the key question of interest.
 - INCENTIVES vs NONINCENTIVES vs. control. This is the secondary coefficient of interest for exploring whether the incentivized version of Generasi works better than the non-incentivized version of Generasi. For the academic economics paper, this is the key question of interest.
- For the main regressions, we will show the results in the Wave II data and the Wave III data. We will then calculate and report an “average” Generasi effect pooling the two waves of the program, as described below.
- All regression will control for the baseline value of the dependent variable (average baseline for kecamatan, dummy for having an individual-specific panel baseline value, and that panel baseline value (0 if unavailable), includes Kab FE (since this is the level of stratification). All household – survey regressions include SAMPLE dummies for how the household was sampled (i.e., which of the three sampling categories it was in interacted with whether it was a panel or cross-section household). Regressions are unweighted to maximize power (although summary statistics of means are calculated using weights). We include age dummies for all child variables (health and education). No other covariates will be included.
 - We will report robustness versions of the regressions where we:
 - Use only the average baseline for kecamatan (don’t use individual control)
 - Don’t control for the baseline level or any other covariates
 - Average everything to the kecamatan level and run kecamatan level regressions controlling for the average baseline for the kecamatan

But these are not our main specifications.

- We drop all kecamatans where we know ex-ante they should have been dropped, but will report reduced form on entire 300 sample as a robustness check. The ones to drop are UPP (early list), Spada, and kecerdasan as of October 2006. Since our randomizing predicts treatment almost perfectly (only 1 kecamatan off) after these ex-ante lists are dropped, we can run regressions directly. This is ITT but will be virtually identical to TOT since there is only one kecamatan that is a noncomplier.
- For the analysis of Wave II data, we need to deal with the fact that not all places randomized to receive Generasi received it. In particular:
 - In year 1, all places originally scheduled to have KDP in year 1 and randomized to receive Generasi were funded
 - Conditional on not having KDP in year 1, we held an additional lottery to add some additional kec in year 1. This extra lottery was stratified by province.
 - To account for the stratification in this additional lottery, we also need to include:
 - Province * PreviousKDPEXperienceFE
 - Note that for the final Wave III analysis, we don’t have these complications, so in theory we could just use district FE for simplicity, since we’ll be back to the original randomization variable. However, we will include the province * PreviousKDPEXperience fixed effects to maintain consistency with the Wave II analysis.
- The Wave II survey was fielded during the period when Generasi Year 2 locations had begun planning but had not received any money. These kecamatans will be treated as controls for the analysis of the Wave II data but will be treated as treatment for the analysis of the Wave III data. Thus GENERASI_2 (Generasi at time t=2) is 1 for Year 1 kecamatans and 0 for Year 2 kecamatans and controls; GENERASI_3 (Generasi at time t=3) is 1 for Year 1 and Year 2 kecamatans and 0 for controls.

- For computing the average effect over Wave II and Wave III, all control variables (e.g., district FE, sample controls, baseline values, etc) will be fully interacted with wave dummies, to capture the fact that there may be differential trends in different parts of the country.
- Thus the regressions we will run for each indicator are:
- Generasi vs control:

- Wave II:

$$y_{pdsi2} = \alpha_d + \beta_1 GENERASI_{pds2} + \gamma_1 y_{pdsi1} + \gamma_2 1_{\{ypdsi1 \neq mis\ sin\ g\}} + \gamma_3 \overline{y_{pds1}} + SAMPLE_{pdsi} + \alpha_p \times P_s + \varepsilon_{pdsi}$$

- Wave III:

$$y_{pdsi3} = \alpha_d + \beta_1 GENERASI_{pds3} + \gamma_1 y_{pdsi1} + \gamma_2 1_{\{ypdsi1 \neq mis\ sin\ g\}} + \gamma_3 \overline{y_{pds1}} + SAMPLE_{pdsi} + \alpha_p \times P_s + \varepsilon_{pdsi}$$

- Wave II an Wave III combined average effect

$$y_{pdsit} = \alpha_{dt} + \beta_1 GENERASI_{pdst} + \gamma_{1t} y_{pdsi1} + \gamma_{2t} 1_{\{ypdsi1 \neq mis\ sin\ g\}} + \gamma_{3t} \overline{y_{pds1}} + \pi_t SAMPLE_{pdsi} + \alpha_{pt} \times P_s + \varepsilon_{pdsit}$$

- Generasi vs control, incentives vs. not:

- Wave II:

$$y_{pdsi2} = \alpha_d + \beta_1 GENERASI_{pds2} + \beta_1 GENERASI_INCENTIVES_{pds2} + \gamma_1 y_{pdsi1} + \gamma_2 1_{\{ypdsi1 \neq mis\ sin\ g\}} + \gamma_3 \overline{y_{pds1}} + SAMPLE_{pdsi} + \alpha_p \times P_s + \varepsilon_{pdsi}$$

- Wave III:

$$y_{pdsi3} = \alpha_d + \beta_1 GENERASI_{pds3} + \beta_1 GENERASI_INCENTIVES_{pds3} + \gamma_1 y_{pdsi1} + \gamma_2 1_{\{ypdsi1 \neq mis\ sin\ g\}} + \gamma_3 \overline{y_{pds1}} + SAMPLE_{pdsi} + \alpha_p \times P_s + \varepsilon_{pdsi}$$

- Wave II an Wave III combined average effect

$$y_{pdsit} = \alpha_{dt} + \beta_1 GENERASI_{pdst} + \beta_1 GENERASI_INCENTIVES_{pdst} + \gamma_{1t} y_{pdsi1} + \gamma_{2t} 1_{\{ypdsi1 \neq mis\ sin\ g\}} + \gamma_{3t} \overline{y_{pds1}} + \pi_t SAMPLE_{pdsi} + \alpha_{pt} \times P_s + \varepsilon_{pdsit}$$

where y_{dsit} is the outcome in Wave t , α_d is a kabupaten fixed effect, y_{dsi1} is the baseline value for individual i (assuming that this is a panel household, and 0 if it is not a panel household), $1_{\{ypdsi1 \neq mis\ sin\ g\}}$ is a dummy for being a panel household, $\overline{y_{ds1}}$ is the average baseline value for the kecamatan, SAMPLE are dummies for how the household was sampled interacted with being a panel or cross-section household, and $\alpha_p \times P_s$ are province-specific dummies for being in the previous-KDP sample. Standard errors are clustered at the subdistrict level in all specifications.

- The robustness regressions are:
 - Include only the ybar for the kecamatan, not the panel specific information.
 - Don't control for baseline level at all or any other covariates
 - Aggregate to the kecamatan level controlling for baseline average level.
 - Do full intent-to-treat analysis on the entire sample of 300 (Wave III only)
 - Do full intent-to-treat analysis on the entire sample of 300 with no controls (Wave III only)
- For each family of indicators, we will test average impacts as follows:
 - Within each family we construct the average treatment effects across all indicators in the family. We will run the regressions together (i.e., stacked regressions clustered by village to allow arbitrary variance-covariance matrix within each village) and compute the average affect across all the indicators in the family, where each effect beta is normalized by the standard deviation of the indicator.
 - Note that if a family consists entirely of binary variables, we will use the average effect, rather than the average standardized effect.

- Sample:
 - We drop all 9 kecamatans where “SPADA LIST” = 1. This is a list that was defined before randomization where Generasi was not allowed to take place. This was a communication failure – should not have been included in randomization list.
 - We have an ex-ante list of 20 UPP kecamatans from Prahas. Of the ones on this list, only 1 gets Generasi each year. We will drop all of them.
 - We drop all 8 Kecamatan Bermasalah from October 2006 and do the same thing with that.
- Mixed control issue
 - The control is more likely to get PNPM regular than the treatment groups (which get it with probability 0). Thus our control group is slightly mixed. I don’t think we should do anything about this in the main analysis but it is worth noting this for future reference.

Sample Tables:

**Main analysis tables (shown with variables from Section 1.2 for illustration – follow these two formats for all analysis in section 1,3,4)
Changes over time analysis (Section 2.2) also will follow this format**

Overall Impact

Indicator	Wave II (2008)			Wave III (2009)		AVERAGE
	Baseline mean (1)	Control Mean (2)	Generasi Treatment Effect (3)	Control Mean (4)	Generasi Treatment Effect (5)	Average Treatment Effect (6)
<i>Main 12 indicators</i>						
Number prenatal visits						
Delivery by trained midwife						
Number of postnatal visits						
Iron tablet sachets						
Percent of immunization						
Number of weight checks						
Number Vitamin A supplements						
Percent malnourished						
Age 7–12 gross enrollment						
Age 13–15 gross enrollment						
Age 7–12 gross attendance						
Age 13–15 gross attendance						6.
<i>Additional Education Indicators</i>						
Age 13–15 conditional attendance						
Age 13–15 enrolled in SMP						
Age 7–12 enrolled in SD						
Age 13–15 enrolled other than SMP						

Kecamatan SD gross enrollment

Kecamatan SMP gross enrollment

Average standardized effect

Average standardized effect health

Average standardized effect educ.

A vs. B

Indicator	Wave II			Wave III			AVERAGE		
	Versi A Treatment Effect	Versi B Treatment Effect	Versi A Additional Effect	Versi A Treatment Effect	Versi B Treatment Effect	Versi A Additional Effect	Versi A Average Treatment Effect	Versi B Average Treatment Effect	Versi A Average Additional Effect
<i>Main 12 indicators</i>									
Number prenatal visits									
Delivery by trained midwife									
Number of postnatal visits									
Iron tablet sachets									
Percent of immunization									
Number of weight checks									
Number Vitamin A supplements									
Percent malnourished									
Age 7–12 gross enrollment									
Age 13–15 gross enrollment									
Age 7–12 gross attendance									
Age 13–15 gross attendance									
<i>Additional Education Indicators</i>									
Age 13–15 conditional attendance									
Age 13–15 enrolled in SMP									
Age 7–12 enrolled in SD									
Age 13–15 enrolled other than SMP									
Kecamatan SD gross enrollment									
Kecamatan SMP gross enrollment									
Average standardized effect									

Average standardized effect health

Average standardized effect educ.

Heterogeneity Analysis Sample Table

Indicator	Wave II			Wave III			AVERAGE		
	Generasi Effect	Interaction with Pre-Period Level	Generasi at 10 th Percentile	Generasi Effect	Interaction with Pre-Period Level	Generasi at 10 th Percentile	Generasi Effect	Interaction with Pre-Period Level	Generasi at 10 th Percentile
<i>Main 12 indicators</i>									
Number prenatal visits									
Delivery by trained midwife									
Number of postnatal visits									
Iron tablet sachets									
Percent of immunization									
Number of weight checks									
Number Vitamin A supplements									
Percent malnourished									
Age 7–12 gross enrollment									
Age 13–15 gross enrollment									
Age 7–12 gross attendance									
Age 13–15 gross attendance									
<i>Additional Education Indicators</i>									
Age 13–15 conditional attendance	7.	8.	9.	10.	11.	12.	13.	14.	15.
Age 13–15 enrolled in SMP	18.	19.	20.	21.	22.	23.	24.	25.	26.
Age 7–12 enrolled in SD									
Age 13–15 enrolled other than SMP									
Kecamatan SD gross enrollment									

Kecamatan SMP gross enrollment

Average standardized effect

Average standardized effect health

Average standardized effect educ.

- Notes
 - MIS data
 - We need:
 - Budgets
 - Susan to check on to make sure we have Year I and Year II data at least, if not Year III
 - Susan to remind Gerda to push to get complete budget
 - Sasaran
 - Handicaps
 - Scores
 - Susan to follow up on MIS data
 - This will be used for Section 4.1 and 4.4 above
- Non-users: no separate analysis required, but we will just note this
 - 10% of funds for non-users in Java in Year 2
 - 25% of funds for non-users off Java in Year 3
- Majalenka:
 - We will include Majalengka as treated, so no adjustments were done
- Academic paper other notes:
 - For the “corruption” section, we can interact the “how well do you know the kepala desa/elites” question with the direct benefits to see if the incentives reduce targeting towards elites
- Test scores
 - Susan to follow up with Yulia as to who is being tested. Also do we know what the test rate is?
- Timetable and workplan
 - Jan 15: Finalize the analysis document
 - Feb 1: Joey can start working on analysis and making new tables
 - Joey should cc: all questions to Ben and Junko since Ben many not be able to respond
 - Apr 15: Get full cleaned data
 - May 30: Have basic analysis complete based on analysis plan.
 - June 1-June 30: Write WB report.
 - May 15-August 31: Ben write academic paper
 - October: presentation in Jakarta
- Who does what?
 - Joey works full time Feb 1 -> March 22
 - Joey and Lina work full time March 22 -> April 16 (split up remaining work between Joey and Lina)
 - April 16 ->May 30: Joey for sure, Lina tentative
 - After Joey leaves Lina does incidental analysis