Dynamic Document for Fiscal Impacts of Deworming

#starting point is  
# Calcs-Table 5!N21 =  
# 'Assumps&Panel A Calcs'!$B$135 \* 'Assumps&Panel A Calcs'!$B$10 \*   
# 'Model Params&Exp Profiles'!R9 \* 'Model Params&Exp Profiles'!R8

## Data   
gov\_bonds <- 0.1185 #Kenyan interest on sovereign debt - Central Bank of Kenya  
inflation <- 0.02 #Kenyan inflation rate - World Bank Development Indicators  
wage\_ag\_val <- 11.84 #Mean hourly wage rate (KSH) - Suri 2011   
wage\_ww\_val <- 14.5850933 #Control group hourly wage, ww (cond >=10 hrs per week) - Table 4, Panel B  
profits\_se\_val <- 1766 #Control group monthly self-employed profits - Table 4, Panel A  
hours\_se\_cond\_val <- 38.1 #Control group weekly self-employed hours, conditional on hrs >0 - Table D13, Panel D  
hours\_ag\_val <- 8.3 #Control group hrs per week, agriculture - Table 4, Panel D  
hours\_ww\_val <- 6.9 #Control group hrs per week, working for wages - Table 4, Panel B  
hours\_se\_val <- 3.3 #Control group hrs per week, self-employment - Table 4, Panel A  
ex\_rate\_val <- 85 #Exchange Rate - Central Bank of Kenya   
growth\_rate\_val <- 1.52/100 #Per-capita GDP growth, 2002-2011 (accessed 1/29/13) - World Bank - see notes  
coverage\_val <- 0.681333333 #Fraction of treated primary school students within 6 km - from W@W - see note  
saturation\_val <- 0.511 #Overall Saturation - not reported in table, average of T & C  
full\_saturation\_val <- 0.75 #REPEATED WITH q\_full\_val in Research component  
tax\_val <- 0.16575 #ADD INFO  
unit\_cost\_local\_val <- 43.66 #Deworm the World  
years\_of\_treat\_val <- 2.41 #Additional Years of Treatment - Table 1, Panel A  
   
   
## Research   
lambda1\_vals <- c(3.49, 0) #Hrs per week increase for men and women CONFIRM  
lambda2\_val <- 10.2 #Externality effect (proportional) - Table 3, Panel B  
q\_full\_val <- 0.75 #Take up rates with full subsidy. From Miguel and Kremmer (2007)  
   
## Guess work   
periods\_val <- 50 #Total number of periods to forecast wages  
time\_to\_jm\_val <- 10 #Time from intial period until individual join the labor force  
coef\_exp\_val <- c(0, 0) #Years of experience coefficients (1-linear, 2-cuadratic) - see notes   
teach\_sal\_val <- 5041 #Yearly secondary schooling compensation 5041 - from ROI materials  
teach\_ben\_val <- 217.47 #Yearly secondary schooling teacher benefits 217.47  
n\_students\_val <- 45 #Average pupils per teacher 45  
  
  
# (Delta E) Additional direct seconday schooling increase (from Joan)   
delta\_ed\_vals <- c(-0.00176350949079451, 0.00696052250263997, 0.0258570306763183,   
 0.0239963665555466, 0.027301406306074, 0.0234125454594173, 0.0279278879439199,   
 0.00647044449446303, 0.00835739437790601)   
delta\_ed\_vals <- cbind(delta\_ed\_vals, 1999:2007)  
  
#Additional externality secondary schooling increase (from Joan)  
delta\_ed\_ext\_vals <- c(-0.0110126908021048, 0.0140448546741008, -0.0034636291545585,  
 0.0112940214439477, 0.0571608179771775, -0.0560546793186931,  
 0.0558284756343451, 0.1546264843901160, 0.0055961489945619)  
delta\_ed\_ext\_vals <- cbind(delta\_ed\_ext\_vals, 1999:2007)  
  
#Notes:   
# on growth\_rate\_val: (http://data.worldbank.org/indicator/NY.GDP.PCAP.KD/), see calculation on "Kenya GDP per capita" tab. In W@W this equals 1.52%. ISSUE: This growth number should be updated to be 2002-2014, I think.  
# on coef\_exp\_val: 1998/1999 Kenyan labor force survey; regression of earnings on age, age^2, female dummy, indicators for attained primary/secondary/beyond, and province dummies. Estimate used in W@W: (0.1019575, -0.0010413). ISSUE: For now assume no further life cycle adjustment beyond KLPS-3 (likely a conservative assumption).  
# coverage\_val: Overall Saturation (0.511) / 0.75 - not reported in table, average of T & C

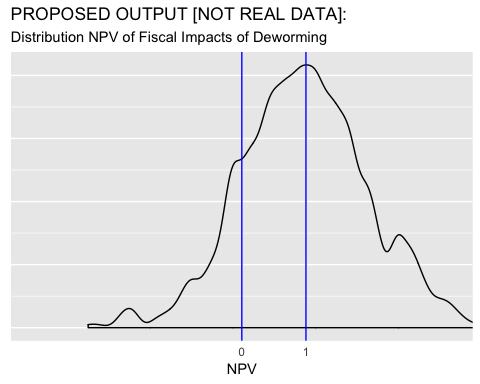
## Who are the policy makers?

* Ministries of Health, Education and Finance.

## What is the relevant output that should be used to inform them?

* NPV

## Proposed output:



## Main Equation (the model)

# Gamma is used to index gender.   
npv <- function(n\_male=1/2, n\_female=1/2,   
 interest\_r=interst\_r\_val,   
 wage=wage\_t\_val,   
 lambda1\_male=lambda1\_vals[1],  
 lambda1\_female=lambda1\_vals[2],   
 tax=tax\_val,   
 saturation=saturation\_val,   
 coverage=coverage\_val,   
 cost\_of\_schooling=cost\_per\_student,   
 delta\_ed\_male=delta\_ed\_vals[,1],   
 delta\_ed\_female=delta\_ed\_vals[,1],   
 lambda2\_male=lambda2\_vals[1],   
 lambda2\_female=lambda2\_vals[2],   
 s1=0, q1=0, s2=s2\_val, q2=q2\_val,   
 periods=periods\_val) {  
 ns <- c(n\_male, n\_female)  
 lambda1s <- c(lambda1\_male, lambda1\_female)  
 lambda2s <- c(lambda2\_male, lambda2\_female)  
 index\_t <- 0:periods  
 delta\_ed\_s <- cbind(delta\_ed\_male, delta\_ed\_female)   
 delta\_ed\_s <- rbind(c(0,0), delta\_ed\_s, matrix(0,41, 2) )   
   
 benef <- matrix(NA, 51,2)  
 for (i in 1:2){  
 benef[,i] <- ( 1 / (1 + interest\_r) )^index\_t \* wage \*   
 ( lambda1s[i] + saturation \* lambda2s[i] / coverage )   
 }  
   
 res1 <- sum( ns \* ( tax \* apply(benef, 2, sum) -   
 apply( ( 1 / (1 + interest\_r) )^index\_t \*   
 delta\_ed\_s \* cost\_of\_schooling, 2, sum) )  
 ) - (s2 \* q2 - s1 \* q1)   
# browser()  
 return(res1)   
}

## Sub components:

### 1 - “”

The real interest rate is obtained from the interest rate on goberment bonds (0.118) minus the inflation rate (0.02).

interst\_r\_val <- gov\_bonds - inflation

### 2 - “”

ndividual in the data are assumed to enter the labor force 10 years after the (data) present day ( for ). Wage at time is the weekely starting wage in USD () that has a base growth rate equal to the per capita GDP growth () applied to however many years of work (). In addition to this growth, the salaries are adjusted to represent a (concave) wage life cycle profile ().

#### 2.1 - “”

The initial wage in dollars () is a weighted average of wages for control group in agriculture, working wage, and self-employed sectors (). The weights correspond to the average number of hours in each sector () relative to the sum of the average number of hours in each sector.

The wage in agriculture comes from research (Suri, 2011), the working wage comes from the data and its defined as hourly wage for the control group for those who reported more than 10 hrs of work per week. The self-employed wage () was constructed as follows:

Where both parameters (Monthly self-employed profits and self-employed hours for the control group, conditional on hrs >0 - -) come from the data (ww paper).

* **Question**: hrs ag = Assumps&Panel A Calcs B25 - still not sure how it came from table 4 (no panel D and )
* **Question**: explain diff betwee and (why used differently)

wage\_0\_f <- function(wage\_ag = wage\_ag\_val,   
 wage\_ww = wage\_ww\_val,   
 profits\_se = profits\_se\_val,   
 hours\_se\_cond = hours\_se\_cond\_val,   
 hours\_ag = hours\_ag\_val,   
 hours\_ww = hours\_ww\_val,   
 hours\_se = hours\_se\_val,   
 ex\_rate = ex\_rate\_val) {  
 wage\_se <- profits\_se / (4.5 \* hours\_se\_cond)  
 wage\_ls <- c(wage\_ag, wage\_ww, wage\_se)  
 alpha\_ls <- c(hours\_ag, hours\_ww, hours\_se) / sum( c(hours\_ag, hours\_ww, hours\_se) )  
 res1 <- 1/ex\_rate \* sum( wage\_ls \* alpha\_ls )  
 return(res1)  
}  
  
#close to value from spreadsheet (Assumps&Panel A Calcs!B137 = 0.1481084),   
#but I suspect diff due to computational precision   
  
wage\_0\_val <- wage\_0\_f()   
  
experience\_val <- 0:periods\_val - time\_to\_jm\_val  
  
wage\_t <- function(wage\_0 = wage\_0\_val,   
 growth\_rate = growth\_rate\_val,   
 experience = experience\_val,   
 coef\_exp1 = coef\_exp\_val[1],   
 coef\_exp2 = coef\_exp\_val[2]) {  
 res1 <- 52 \* wage\_0 \*( ( 1 + growth\_rate )^experience ) \*   
 ( 1 + coef\_exp1 \* experience + coef\_exp2 \* experience^2 ) \*   
 ifelse(0:periods\_val >= time\_to\_jm\_val, 1, 0)  
 return(res1)   
}  
  
#close to value from spreadsheet (Calcs-Table 5!N21.. = 7.701634678),   
#but I suspect diff due to computational precision   
wage\_t\_val <- wage\_t()

### 3 - “” and “”

represents the estimated impact of deworming on hours of work for men a womam. This two parameter are combined with a unwweighted mean:

Its components come from research (W@W).

the estimated externality effect (EXPLAIN) and comes from research (W@W). Note that this parameter in not estimated by gender, so we repeat its value two times.

lambda1\_vals <- rep(0.5 \* lambda1\_vals[1] + 0.5 \*lambda1\_vals[2], 2)  
lambda2\_vals <- rep(lambda2\_val, 2)

### 4 - and

is the fraction of treated primary school students within 6 km. Computed as

is the saturation of the intervention, defined as:

**Note:** there is a circularity on how parameters are defined here. Everything depends on Full Saturation () at the end of the day.

#R   
coverage\_val <- saturation\_val / full\_saturation\_val  
#p   
saturation\_val <- full\_saturation\_val \* coverage\_val

### 5 - and

represents the cost per student. This is calculated as the salary of the teacher plus benefits, divided by the average number of students per teacher.

cost\_per\_student <- (teach\_sal\_val + teach\_ben\_val) / n\_students\_val

For we use a series of estimated effects the additional direct increase in seconday schooling from 1999 to 2007 obtained from [need to define the source “from Joan” in Assumps&Panel A Calcs!A93].

This series does not take into account the externality effects. To incorporate the we need another series (same source) that estimates the additional secondary schooling increase due to the externality and add it to the original series.

# Nothing here yet with delta\_ed\_vals, but would like to incorporate model from Joan  
# delta\_ed\_vals <-   
delta\_ed\_ext\_total <- delta\_ed\_ext\_vals[,1] + delta\_ed\_vals[,1]

**Note:** need to understand better the date of each component (of the model, not only this section).

### 6 -

#### 6.1 -

There is no subsidy for deworming under the status quo.

#### 6.2 - : complete subsidy to per capita costs of deworming.

With complete subsidy, represents the total direct costs of deworming in USD. Calculated as follows

#### 6.3 -

The take-up with full subsidy () comes from a previous study (Miguel and Kremer 2007) and takes the value of 0.75.

s2\_val <- ( unit\_cost\_local\_val / ex\_rate\_val ) \* years\_of\_treat\_val  
q2\_val <- q\_full\_val

### Main results

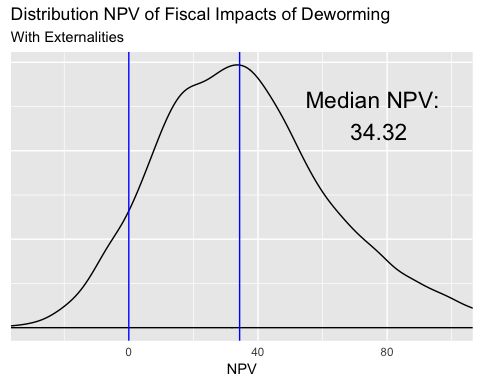
#no externality NPV  
res\_npv\_no\_ext <- npv(lambda2\_male = 0, lambda2\_female = 0)  
  
#yes externality NPV  
res\_npv\_yes\_ext <- npv(delta\_ed\_male = delta\_ed\_ext\_total,   
 delta\_ed\_female = delta\_ed\_ext\_total )

* **NPV without externalities ():** -0.6097
* **NPV with externalities ( ):** 34.3187

## Montecarlo simulations

Describe approach to MC: - Eveything normal with sd= 0.1 \* mean  
- Need to work with experts to add more realistic parameters

#34.8401  
nsims <- 1e4  
set.seed(1234)  
#Defaoult dist: normal, default sd: 0.1\* mean  
## Data   
gov\_bonds\_sim <- rnorm(n = nsims, mean = gov\_bonds, sd = 0.1 \* gov\_bonds)   
inflation\_sim <- rnorm(nsims, inflation, 0.1 \* inflation)  
wage\_ag\_val\_sim <- rnorm(nsims, wage\_ag\_val, 0.1 \* wage\_ag\_val)  
wage\_ww\_val\_sim <- rnorm(nsims, wage\_ww\_val, 0.1 \* wage\_ww\_val)  
profits\_se\_val\_sim <- rnorm(nsims, profits\_se\_val, 0.1 \* profits\_se\_val)  
hours\_se\_cond\_val\_sim <- rnorm(nsims, hours\_se\_cond\_val, 0.1 \* hours\_se\_cond\_val)  
hours\_ag\_val\_sim <- rnorm(nsims, hours\_ag\_val, 0.1 \* hours\_ag\_val)  
hours\_ww\_val\_sim <- rnorm(nsims, hours\_ww\_val, 0.1 \* hours\_ww\_val)  
hours\_se\_val\_sim <- rnorm(nsims, hours\_se\_val, 0.1 \* hours\_se\_val)  
ex\_rate\_val\_sim <- rnorm(nsims, ex\_rate\_val, 0.1 \* ex\_rate\_val)  
growth\_rate\_val\_sim <- rnorm(nsims, growth\_rate\_val, 0.1 \* growth\_rate\_val)  
coverage\_val\_sim <- rnorm(nsims, coverage\_val, 0.1 \* coverage\_val)  
saturation\_val\_sim <- rnorm(nsims, saturation\_val, 0.1 \* saturation\_val)  
full\_saturation\_val\_sim <- rnorm(nsims, full\_saturation\_val, 0.1 \* full\_saturation\_val)  
tax\_val\_sim <- rnorm(nsims, tax\_val, 0.1 \* tax\_val)  
unit\_cost\_local\_val\_sim <- rnorm(nsims, unit\_cost\_local\_val, 0.1 \* unit\_cost\_local\_val)  
years\_of\_treat\_val\_sim <- rnorm(nsims, years\_of\_treat\_val, 0.1 \* years\_of\_treat\_val)  
  
## Research  
lambda1\_vals\_sim <- sapply(lambda1\_vals, function(x) rnorm(nsims, mean = x, sd = 0.1 \* x) )  
lambda2\_val\_sim <- rnorm(nsims, lambda2\_val, 0.1 \* lambda2\_val)  
q\_full\_val\_sim <- rnorm(nsims, q\_full\_val, 0.1 \* q\_full\_val)  
  
## Guess work  
periods\_val <- 50 #Total number of periods to forecast wages  
time\_to\_jm\_val <- 10 #Time from intial period until individual join the labor force  
coef\_exp\_val\_sim <- sapply(coef\_exp\_val, function(x) rnorm(nsims, mean = x, sd = 0.001))  
teach\_sal\_val\_sim <- rnorm(nsims, teach\_sal\_val, 0.1 \* teach\_sal\_val)  
teach\_ben\_val\_sim <- rnorm(nsims, teach\_ben\_val, 0.1 \* teach\_ben\_val)  
n\_students\_val\_sim <- rnorm(nsims, n\_students\_val, 0.1 \* n\_students\_val)  
  
delta\_ed\_vals\_sim <- sapply(delta\_ed\_vals[,1], function(x) rnorm(nsims, mean = x, sd = sd(delta\_ed\_vals[,1])))  
colnames(delta\_ed\_vals\_sim) <- 1999:2007  
  
delta\_ed\_ext\_vals\_sim <- sapply(delta\_ed\_ext\_vals[,1], function(x) rnorm(nsims, mean = x, sd = sd(delta\_ed\_ext\_vals[,1])))  
colnames(delta\_ed\_ext\_vals\_sim) <- 1999:2007  
  
npv\_sim <- rep(NA, nsims)  
#yes externality NPV  
for (i in 1:nsims) {  
 interst\_r\_val <- gov\_bonds\_sim[i] - inflation\_sim[i]  
 wage\_0\_val <- wage\_0\_f(wage\_ag = wage\_ag\_val\_sim[i],   
 wage\_ww = wage\_ww\_val\_sim[i],   
 profits\_se = profits\_se\_val\_sim[i],   
 hours\_se\_cond = hours\_se\_cond\_val\_sim[i],   
 hours\_ag = hours\_ag\_val\_sim[i],   
 hours\_ww = hours\_ww\_val\_sim[i],   
 hours\_se = hours\_se\_val\_sim[i],   
 ex\_rate = ex\_rate\_val\_sim[i])   
 experience\_val <- 0:periods\_val - time\_to\_jm\_val  
 wage\_t\_val <- wage\_t(wage\_0 = wage\_0\_val,   
 growth\_rate = growth\_rate\_val\_sim[i],   
 experience = experience\_val,   
 coef\_exp1 = coef\_exp\_val\_sim[i,1],   
 coef\_exp2 = coef\_exp\_val\_sim[i,2])  
 lambda1\_vals\_aux <- rep(0.5 \* lambda1\_vals\_sim[i,1] + 0.5 \* lambda1\_vals\_sim[i,2], 2)  
 lambda2\_vals <- rep(lambda2\_val\_sim[i], 2)  
 coverage\_val\_aux <- saturation\_val\_sim[i] / full\_saturation\_val\_sim[i]  
 saturation\_val\_aux <- full\_saturation\_val\_sim[i] \* coverage\_val\_sim[i]  
 cost\_per\_student <- (teach\_sal\_val\_sim[i] + teach\_ben\_val\_sim[i]) / n\_students\_val\_sim[i]  
 q2\_val\_aux <- q\_full\_val\_sim[i]  
 s2\_val\_aux <- ( unit\_cost\_local\_val\_sim[i] / ex\_rate\_val\_sim[i] ) \* years\_of\_treat\_val\_sim[i]  
 delta\_ed\_ext\_total\_sim <- delta\_ed\_vals\_sim[i,] + delta\_ed\_ext\_vals\_sim[i,]  
   
 include\_ext <- TRUE  
 if (include\_ext==TRUE){  
 delta\_ed\_final <- delta\_ed\_ext\_total\_sim  
 }else{  
 delta\_ed\_final <- delta\_ed\_vals\_sim[i,]  
 }  
   
 npv\_sim[i] <- npv(interest\_r = interst\_r\_val,   
 wage = wage\_t\_val,   
 lambda1\_male = lambda1\_vals\_aux[1],   
 lambda1\_female = lambda1\_vals\_aux[2],   
 lambda2\_male = lambda2\_vals[1],   
 lambda2\_female = lambda2\_vals[2],  
 coverage = coverage\_val\_aux,  
 saturation = saturation\_val\_aux,  
 tax = tax\_val\_sim[i],   
 cost\_of\_schooling=cost\_per\_student,   
 delta\_ed\_male = delta\_ed\_final,   
 delta\_ed\_female = delta\_ed\_final,   
 q2 = q2\_val\_aux,   
 s2 = s2\_val\_aux)  
}  
  
# unit test  
if (abs(sd(npv\_sim) - 30.5607)>0.0001 ) {  
 print("Output has change")  
}  
  
npv\_for\_text <- paste("Median NPV:\n ", round(median(npv\_sim), 2))  
ggplot() +  
 geom\_density(aes(x = npv\_sim,  
 alpha = 1/2), kernel = "gau") +  
 geom\_vline(xintercept = c(0, median(npv\_sim)), col="blue") +  
 coord\_cartesian(xlim = c(-30,100)) +  
 guides(alpha = "none", colour="none") +  
 labs(y = NULL,  
 x = "NPV" ,  
 title = "Distribution NPV of Fiscal Impacts of Deworming",   
 subtitle = "With Externalities")+  
 annotate("text", x = 2.2 \* median(npv\_sim), y = 0.012, label = npv\_for\_text, size = 6)+  
 theme(axis.ticks = element\_blank(), axis.text.y = element\_blank())



## Sensitivity Analysis

Describe how we can move each component. Link to shiny app.

If we combinde all the subcomponents of equation 1 into on single, very large, expression, we obtain:

## Additional documentation

## Questions (for Michael, Ted or Grace)

* where are the in the spreadsheet?
* To me: should I add small output after each subsection?
* What are the terms behind the calculation of ?
* Clarify circularity behing and
* Why is not used for separate groups (men and female)

## Next steps:

* set up each primitive parama as normal(value,sd = something). DONE
* run 1000? 10000? times and obtain figure 1. DONE
* Finish description of last component. DONE
* deploy shiny app with one slider. DONE
  + wrap everything in a horrible function DONE
  + add aditional sliders DONE
* Display value for NPV
* reproduce results with and without externalities reactively