# ESM 228 Practicum 2

Simone Albuquerque, Sara Orofino, Elena Zhang
May 4, 2020

---

## Theory of Change

**Organization:** American Rivers
**Program Title:** American Rivers Fuel Reduction Program
**Region:** Sierra Nevada region, specificallly the following four watersheds - American River Watershed, Bear River Watershed, Cosumnes River Watershed, Yuba River Watershed (CABY watersheds)

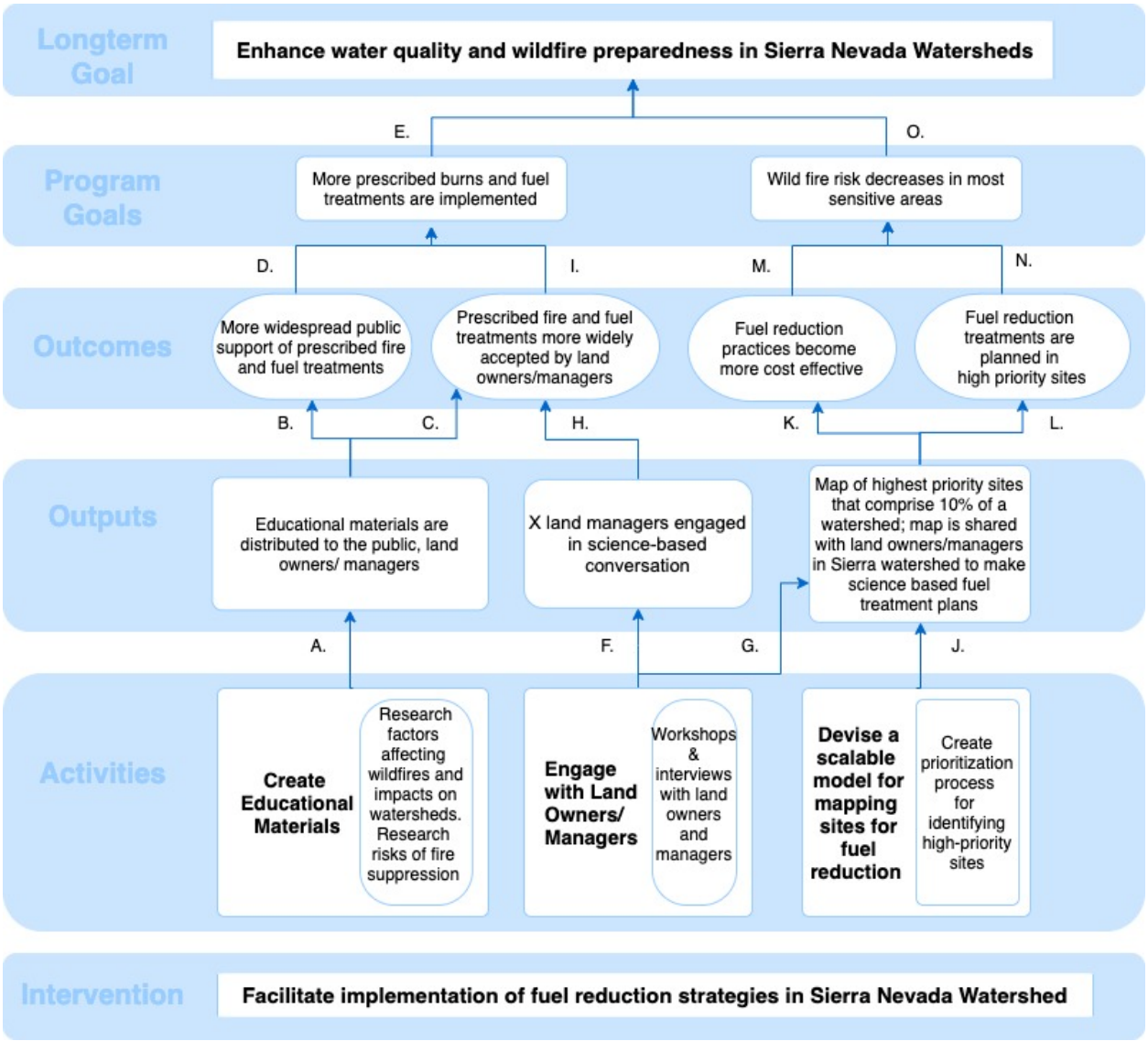


Figure 1: Theory of Change American Rivers Fuel Reduction Program

## Measure Definition

Define the eight most important measures for the Theory of Change.

**Measure 1: Mapping Accuracy**

| | |
|---|---|
| Measure Goal | Assess the accuracy of the mapping of high priority sites |
| Exact Measure | Water quality index before and after wildfires at randomly selected rivers throughout the study area |
| Unit of Measurement | Water quality index consists of the following metrics: temperature (Celsius), turbidity (NTU), bacteria concentration (colony forming units per mL), pH, dissolved oxygen (mg/L) nutrient composition (parts per million), total suspended solids (mg/L), and toxic substances (toxic units) |
| Source of Data | Water quality sampling at regular intervals during wildfire season |
| Rational for Measure | Water quality is the primary factor defining the high priority areas, comparing water quality between randomly sampled sites designated as high priority and those not designated as high priority can help assess how much of the watershed is high priority. |
| Data Collection | American Rivers group |
| Frequency of Collection | Annual |

**Measure 2: Implementation of Fuel Reduction Programs**

| | |
|---|---|
| Measure Goal | Determine if there is a change in the number of fuel reduction programs being implemented on private lands by land owners |
| Exact Measure | Survey question "Do you have a current Cal Fire permit for debris burning?" |
| Unit of Measurement | Individual respondents |
| Source of Data | Independent Survey |
| Rational for Measure | Burning on private property in California requires a permit from Cal Fire, tracking the change in permits over time can give an idea of whether or not the program is increasing the demand for fire permits, which are required to implement fuel reduction programs. There are a lot of external factors (i.e. weather, air quality) that control whether or not burning can actually happen on any given day, but by applying for permits land owners show intent to participate in such programs as other factors allow. |
| Data Collection | Independent survyor contracted by American Rivers |

| | |
|---|---|
| Frequency of Collection | Annually because Cal Fire burn permits are issued annually |

**Measure 3: Change in Wildfire Risk**

| | |
|---|---|
| Measure Goal | Determine the change in frequency of high severity fires in high priority areas over 5 years |
| Exact Measure | Number of fires and total acres burned during each annual fire season in high priority areas |
| Unit of Measurement | Individual wildfire events |
| Source of Data | Cal Fire wildfire statistics |
| Rational for Measure | Implementing a fuel reduction program should reduce wildfire risk. In the event of wildfires there would be less fuel for the fire to burn leading to overall lower acreages burned. |
| Data Collection | Cal Fire collects data on wildfires in California and statistics on the acreage burned per wildfire |
| Frequency of Collection | Annual |

**Measure 4: Affordability of Fuel Reduction**

| | |
|---|---|
| Measure Goal | Determine the number of land owners in high priority sites with willingness to pay |
| Exact Measure | Would you be willing to pay for fuel reduction on your property? Yes or No |
| Unit of Measurement | Individual survey respondents |
| Source of Data | Independent survey firm |
| Rational for Measure | Fuel reduction programs on private land would be paid for by the land owners. In order for fuel programs to be implemented the land owners must be willing to bear the cost of the actions. This measure gives an initial indication of if a land owner would be willing to undertake those costs and allows for tracking how those attitudes change during and after the program. |
| Data Collection | Independent survyor contracted by American Rivers |
| Frequency of Collection | Annual |

**Measure 5: Effectiveness of Educational Materials**

| | |
|---|---|
| Measure Goal | Assess the public perception of prescribed burns and fuel treatments over time |
| Exact Measure | From what you have heard or read, what best describes your opinion of prescribed fires and fuel treatment actions? Oppose, Support, Neither Oppose nor Support (i.e. neutral) |
| Unit of Measurement | Individual survey respondents |
| Source of Data | Independent survey firm |
| Rational for Measure | The best way to know whether the public is supportive of prescribed burns and fuel treatments is to measure it directly. The question does not lead the respondents and has balanced alternatives. The answer "oppose" is placed higher in the list as a signal that it is an appropriate answer. |
| Data Collection | Independent survyor contracted by American Rivers |
| Frequency of Collection | Prior to the educational campaigns and through the program's existence |

**Measure 6: Likelihood of Accepting Prescribed Burns**

| | |
|---|---|
| Measure Goal | Determine whether landowners and managers who engage in science-based conversations are more likely to accept the importance of prescribed fuel treatments. |
| Exact Measure | What is your opinion on the following statement: "Prescribed fire and fuel treatments can lower the risk of wildfire." Strongly disagreee, Disagree, Neutral, Agree, Strongly Agree |
| Unit of Measurement | Individual survey respondent |
| Source of Data | Survey responses from educational meeting attendees in different counties |
| Rational for Measure | The best way to determine the extent to which land managers accept the importance of prescribed fuel treatments is to use a Likert scale question. The response scale starts with "Strongly disagree" and ends with "Strongly Agree", in order to avoid response order bias. A neutral response is added in the middle so respondents aren't forced to choose a side. |
| Data Collection | American Rivers |
| Frequency of Collection | Before and after each educational meeting held as part of the program implementation |

**Measure 7: Impact of Public Support**

| | |
|---|---|
| Measure Goal | Determine if public support increases the number of prescribed burns in the Sierra Nevada region |
| Exact Measure | Satellite data (Modis Terra) imagery of extent of land burned vs. not burned in high priority regions of the study area |
| Unit of Measurement | Acres of study area |
| Source of Data | NASA |
| Rational for Measure | A relatively cost effective means of monitoring large areas for fire and comparing across years |
| Data Collection | NASA collects the Modis Terra satellite data and makes it publically available on their website |
| Frequency of Collection | Readings are taken twice per day, data is analyzed annually during the prescribed burning season |

**Measure 8: Effect on Water Quality**

| | |
|---|---|
| Measure Goal | Determine how water quality changes in the high priority regions of the study area when fuel treatments are used vs. when they are not used. |
| Exact Measure | Sediment yield (total dissolved solids) in reservoir intake areas within 50 miles of a high priority burn site |
| Unit of Measurement | Parts per million |
| Source of Data | Reservoir data |
| Rational for Measure | Reservoirs will already be measuring the water quality in close proximity to infiltration points, which we will target in high priority areas. This is a cost effective way of getting real time water quality data throughout the year, which otherwise may be costly. |
| Data Collection | Local county water utilities |
| Frequency of Collection | Daily, focused on measures during fire season |

## Sampling Strategy

### Measure 2 - Implementation of Fuels Reduction Programs

*Target population*: Private land owners/managers in the study area

Using demogrpahic data of each county, we found the approximate number of homeowners per county in the study region, these values are given in the table below. For this measure, we further assumed 35% of homeowners owned land large enough to implement fuel management practices.

| County | Approximate private homeowner population per county |
|---|---|
| El Dorado | 51900 |
| Placer | 98200 |
| Nevada | 30000 |
| Sierra | 963 |
| Yuba | 15100 |
| Plumas | 6020 |
| Amador | 10900 |
| Alpine | 253 |

*Challenges to representative sampling*: Much of the land in the Sierra Nevada watershed is federally owned not privately owned, but federal actions involve more factors than simply perception and willingness to pay. Private land owners require permits to implement fuel reduction programs but fuel reduction typically doesn't occur every year and it is unclear if the same land owner would continue to apply for permits in years immediately following implementing a fuel reduction program on their land.

*Sampling procedure* Random sampling of private land owners in the study area to determine if they are Cal Fire permit holders, stratified by county since some counties account for a higher proportion of land in the study area.

### DeclareDesign()

Declare the population

```
# baseline = proportion of land area each county contributes to the total study area
# prop = proportion of homeowners in each strata out of the total population
#(workaround for strata mean function error - see 'measure1-delcare-estimator' code below)
# owner = 35% of the number of homeowners in the county, since all homeowners may
#not have large properties where they are participating in fuel management practices


population <- declare_population(
  county = add_level(N=8, # there are 8 counties in the watershed
                     # % coverage of each county in study area
                     baseline=c(0.36,0.23,0.18,0.5,0.5,0.5,0.5, 0.3),
                     prop=c(0.459298084, 0.243157809, 0.141842055, 0.067543836,
                            #proportion of homeowners in each county
```

```
                                0.054035069, 0.028368411, 0.004552455, 0.001202280)),
  # 35% of each county's homeowner population
  owner = add_level(N=c(18000,34000,10500, 337, 5000, 2100, 4000, 89),
                    know=draw_binary(baseline))
)

pop <- population()
pop.vector <- c(18000,34000,10500, 337, 5000, 2100, 4000, 89)

my_estimand <- declare_estimands(mean(know),
                                 label = "Ybar")
```

Reporting and Sampling

```
reporting <- declare_assignment(prob=0.4,
                    assignment_variable = "R")
#lower probably that any given homeowner would be a burn permit holder

sampling <- declare_sampling(strata=county,
                strata_n=c(80,80,80,80,80,80,10,10)) # sample size 500
```

Declare estimator to calculate the strata weighed mean of the sample to estimate the population

```
# function to estimate the population mean using the strata weighted sample mean
strata_weighted_mean <- function(data){
  data.frame(
    estimator_label = "strata_w_mean",
    estimand_label = "Ybar",
    n = nrow(data),
    stringsAsFactors = FALSE,

    estimate = data %>% filter(R==1) %>%
      group_by(as.factor(county)) %>% # strata by county
      summarise(mean=mean(know), prop=mean(prop)) %>% #work around for prop error (see below)
      # mutate(prop=pop.vector/sum(pop.vector)) %>%
      mutate(sub.mean=mean*prop) %>% pull(sub.mean) %>%
      sum())
}
# note: prop column kept throwning me a length error (wanted length 7 not 8),
# I changed all the other vectors to be length 8 and couldn't figure out why it was
# requiring a length 7 when the strata and my pop.vector was set to 8. Patrick helped
# me with a workaround - the prop=mean(prop) gives the proportion of households in the
# strata based on the prop column we defined in the population setup code.
```

Diagnostics

```
answer <- declare_estimator(
  handler = tidy_estimator(strata_weighted_mean),
  estimand = my_estimand)

design <- population + my_estimand + reporting +
          sampling + answer

diagnosis <- diagnose_design(design, sims = 1000)

diagnosis$diagnosands_df[,c(4,5,12,14)] %>%
  kable()
```

| bias | se(bias) | mean_estimate | sd_estimate |
|---|---|---|---|
| 0.0268032 | 0.001355 | 0.3231208 | 0.0463375 |

**Measure 4 - Affordability of Fuel Reduction Programs**

*Target population*: Landowners in high priority areas

*Challenges to representative sampling*: High priority region of the CABY watershed private lands are generally less productive, steeper, and more isolated than non-federal ownerships. Each property is difficult to find and may have a private driveway and/or PO box. This survey will be sent via mail to American River Contacts and all folks in the high priority areas. Since AR contacts have already connected to a coservation organization their response rate will likelu be higher than the individuals who have not. In addition, AR contacts are more likely to be willing to pay for fuels treatments as they have been working with a conservtaion non-profit so they are already selected for interest in the program.

*Sampling procedure*: We will survey homeowners in the High Priority region who have interacted with the fuel management educational program and/or live in the CABY high priority regions and ask whether they would be willing to pay for fuels treatment on their property? a. no b. yes. Sampling wil be conducted by a thir party survey group as residents may be more likely to reposnd yes to American Rivers as they may want to identify with a local consrrvation organizations goals or not admit publicly they would not be willing to pay.

## DeclareDesign()

Declare the population

```
set.seed(228)
population_4 <- declare_population(
  # 3% of homeowners with land large enough for fuel treatment
  households_4 = add_level(N=1400,
      AR_contact=draw_binary(N=N, prob = 0.4),
    WTP=correlate(given = AR_contact, rho = 0.5,
                          draw_binary, prob = 0.5)
))
pop_4 <- population_4()
# using Declare Design here to create a population of 1400 households,
# 500 of are AR contacts. There is a direct correlation between who is
# already in contact with AR and their willingness to pay for fuels treatment

kable(table(pop_4$AR_contact,pop_4$WTP)) %>%
  add_header_above(c("AR Contact"=1,"Consent to do Prescribed Burn"=2))
```

| AR Contact | Consent to do Prescribed Burn | |
|---|---|---|
| | 0 | 1 |
| 0 | 517 | 323 |
| 1 | 182 | 378 |

```
# Generate estimand
my_estimand_4 <- declare_estimands(mean(WTP),
                                   label = "Ybar")
```

Reporting & Sampling

- For residents in contact with AR, the chance of responding is 70%
- For residents not in contact with AR, the chance of responding is 30%

```
# Function for strata weighted mean
strata_weighted_mean_4 <- function(data){
  data.frame(
  estimator_label = "strata_w_mean",
  estimand_label = "Ybar",
  n = nrow(data),
  stringsAsFactors = FALSE,

  estimate = data %>% filter(R==1) %>%
    group_by(AR_contact) %>%
    summarise(mean=mean(WTP)) %>%
    mutate(prop=c(0.3,0.7)) %>%
    mutate(sub.mean=mean*prop) %>% pull(sub.mean) %>%
    sum())
}
```

Diagnostics

```
answer_4 <- declare_estimator(
  handler = tidy_estimator(strata_weighted_mean_4),
  estimand = my_estimand_4)

design_4 <- population_4 + my_estimand_4 + WTP_report +
          WTP_sample + answer_4

diagnosis_4 <- diagnose_design(design, sims = 1000)

diagnosis_4$diagnosands_df[,c(4,5,12,14)] %>%
  kable()
```

| bias | se(bias) | mean_estimate | sd_estimate |
|---|---|---|---|
| 0.028252 | 0.0014891 | 0.3245496 | 0.0460312 |

**Measure 5 - Effectiveness of Educational Materials**

*Target Population*: Private land owners in the studied area

*Challenge of drawing a representative sample*: Sampling efforts are done by mailing questionnaires to randomly selected home owners in the 8 counties in the CABY region, but these clusters are of unequal size. It is unclear how many of the questionnaire recipients actually respond to the questionnaire.

*Sampling procedure*: Random sampling of private land owners in the 8 counties in the CABY region with unequal effort according to potential sample size (since some counties have larger population than others)

Using demogrpahic data of each county, we found the approximate number of private land owners per county in the CABY region.

| County | Approximate private homeowner population per county |
|---|---|
| Placer | 98200 |
| El Dorado | 51900 |
| Nevada | 30000 |
| Yuba | 15100 |
| Amador | 10900 |
| Plumas | 6020 |
| Sierra | 963 |
| Alpine | 253 |

## DeclareDesign()

Declare the population

```
set.seed(228)
population <- declare_population(
  county = add_level(N=8,
           basemean=c(1,0.5,0.5,0,-0.4,-0.2,-1,-1)),
# the 'basemean' variable represents the mean value of the normal random variable
# that we use to generate the scale of the survey answers
# (support, oppose, neither oppose nor support), we assume that land owners in smaller
# counties tend to oppose prescribed burns and fuel treatments.

  homeowner = add_level(N=c(98200,51900,30000,15100,10900,6020,963,253),
                        support=draw_ordered(x = rnorm (n = N, mean = basemean),
                        breaks = c (-1, 1),
                        #break_labels = c("support","neutral","oppose")
                        ))
# Here homeowners are labeled by support (value=3), neutral (value=2), or oppose (value=1),
# corresponding to the value of the normal random variable x<-1, -1<x<1, 1<x
)
pop <- population()
pop.vector <- c(98200,51900,30000,15100,10900,6020,963,253)
```

```
my_estimand <- declare_estimands(mean(support),
                                 label = "Ybar")
```

Reporting and Sampling

```
reporting <- declare_assignment(prob=0.4,
                  assignment_variable = "R")
# we assume a 40% survey response rate.

sampling <- declare_sampling(strata=county,
                strata_n=c(500,200,100,70,70,30,15,15))
# we are sampling by county, our strata. We're going to send out 1000 questionnaires
# in total: 500 questionnaires in Placer, 200 in El Dorado, 100 in Nevada, 70 in Yuba,
# 70 in Amador, 30 in Plumas, 15 in Sierra, and 15 in Alpine.
```

Declare estimator

```
strata_weighted_mean <- function(data){
  data.frame(
  estimator_label = "strata_w_mean",
  estimand_label = "Ybar",
  n = nrow(data),
  stringsAsFactors = FALSE,

  estimate = data %>% filter(R==1) %>%
    group_by(county) %>%
    summarise(mean=mean(support)) %>%
    mutate(prop=pop.vector/sum(pop.vector)) %>%
    mutate(sub.mean=mean*prop) %>% pull(sub.mean) %>%
    sum())
}
```

Diagnostics

```
answer <- declare_estimator(
  handler = tidy_estimator(strata_weighted_mean),
  estimand = my_estimand)

design <- population + my_estimand + reporting + sampling + answer
diagnosis <- diagnose_design(design, sims = 1000)

diagnosis$diagnosands_df[,c(4,5,12,14)] %>%
  kable()
```

| bias | se(bias) | mean_estimate | sd_estimate |
|---|---|---|---|
| -0.000363 | 0.0009626 | 2.296737 | 0.0282978 |