City Graphs

Dina Sinclair

July 17, 2018

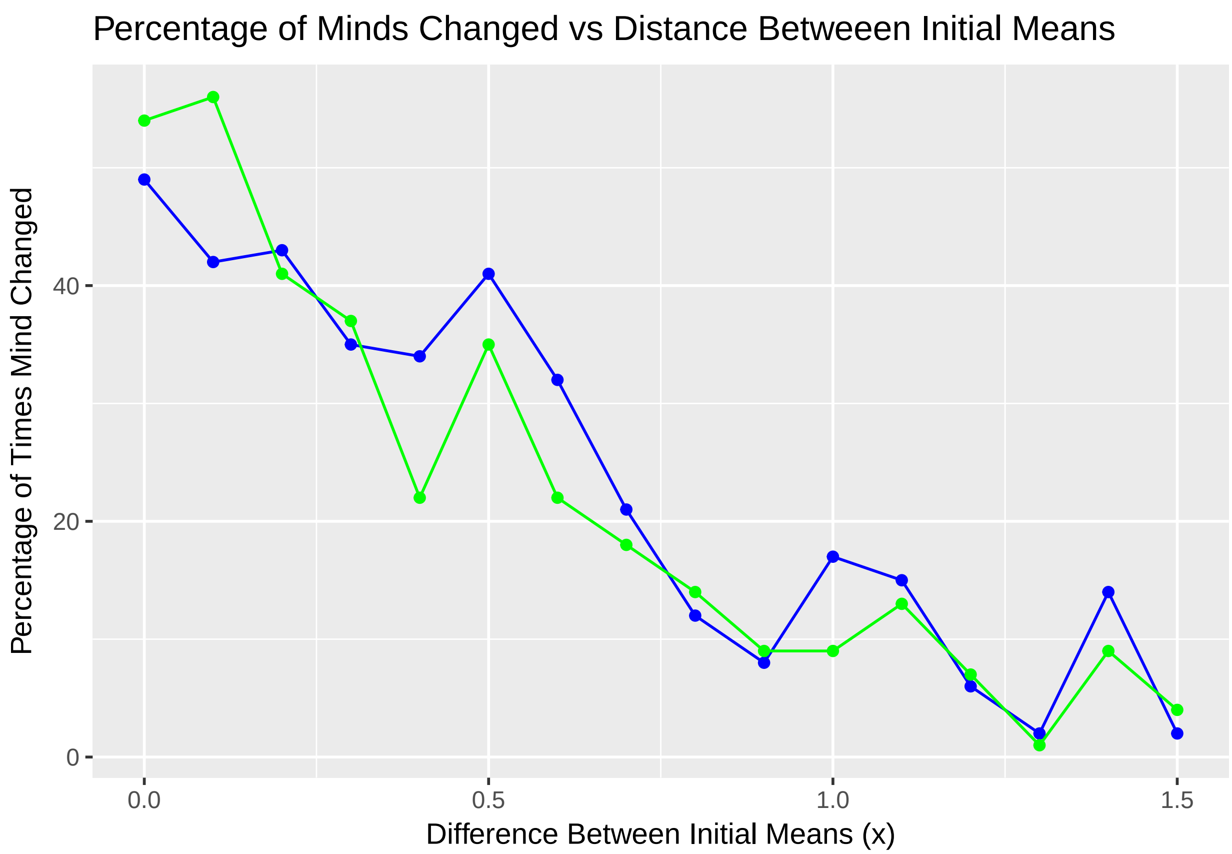
```r
source("C:/Users/Dina/Documents/Oxford/Evaluating Prior Impact/cities_main_code.R")
```

```
## Loading required package: ggplot2

## Loading required package: StanHeaders

## rstan (Version 2.17.3, GitRev: 2e1f913d3ca3)

## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)
```

```r
library(reshape2)
library(ggplot2)
set.seed(17)
```

```r
N <- 16
pilot_c1 <- numeric(N)
pilot_c2 <- numeric(N)
record_i <- numeric(N)
sample <- seq(0,1.5,length.out = N)

for (i in 1:N){
  data <- list(I=2,Y=c(sample[i],0),sigmaSq=c(1,1))
  results <- overall(data=data,num_pilots=1,num_final_cities=1,num_draws=100)
  pilot_c1[i] <- results$nmc[1]
  pilot_c2[i] <- results$nmc[2]
  print(i)
  print(sample[i])

}

df <- data.frame(sample, pilot_c1, pilot_c2)
save(df, file="testing_mean.Rda")
```

```r
load("testing_mean.Rda")
ggplot(df, aes(sample)) +
  geom_line(aes(y=pilot_c1), colour="blue") + geom_point(aes(y=pilot_c1), colour="blue")  +
  geom_line(aes(y=pilot_c2), colour="green") + geom_point(aes(y=pilot_c2), colour="green")  +
  ylab("Percentage of Times Mind Changed") + xlab("Difference Between Initial Means (x)") +
  labs(title="Percentage of Minds Changed vs Distance Betweeen Initial Means")
```

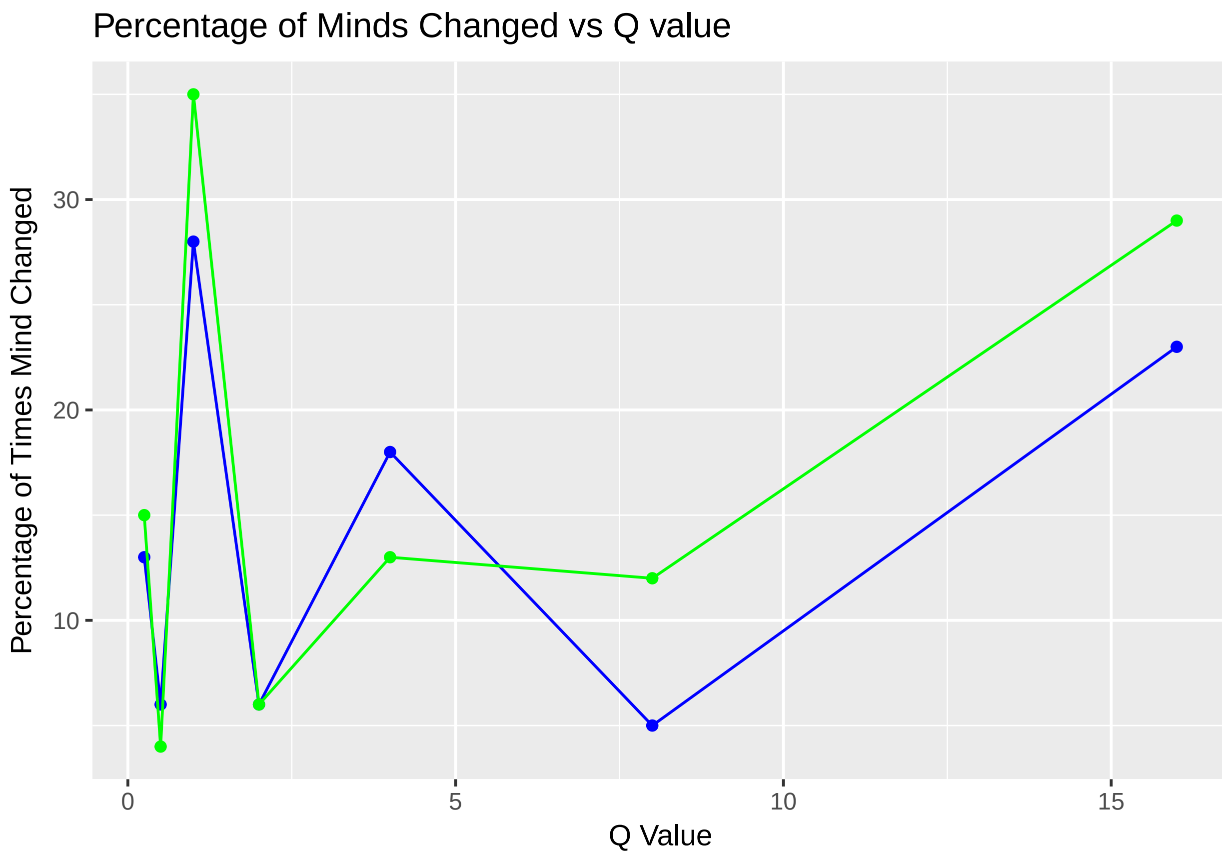


Here, we've run a case where there are only two cities to chose from, we can run a pilot in only one city, and we can run the final program in only one city. City 1 has an initial prior $Y_1 \sim (x, 1)$ where x ranges from 0 to 1.5 and city 2 has an initial prior $Y_2 \sim (0, 1)$. Each simulation has been run 100 times, and the number of times changing the initial ranking (deciding to go for city 2 rather than city 1) has been plotted as the percentage of times we changed our minds (y). Piloting city 1 results plotted in blue, piloting city 2 results plotted in green. As expected, as the means grow farther apart, it becomes less and less likely that a pilot study will change our minds.

```
Q_val <- c(0.25,0.5,1,2,4,8,16)
N <- length(Q_val)
pilot_c1 <- numeric(N)
pilot_c2 <- numeric(N)

for (i in 1:N){
  data <- list(I=2,Y=c(1,0),sigmaSq=c(1,1))
  results <- overall(data=data,num_pilots=1,num_final_cities=1,num_draws=100,Q=Q_val[i])
  pilot_c1[i] <- results$nmc[1]
  pilot_c2[i] <- results$nmc[2]
  print(sample[i])
}

df <- data.frame(Q_val, pilot_c1, pilot_c2)
save(df, file="testing_Q.Rda")
df
```

```
load("testing_Q.Rda")
ggplot(df, aes(Q_val)) +
```

```
  geom_line(aes(y=pilot_c1), colour="blue") + geom_point(aes(y=pilot_c1), colour="blue")  +
  geom_line(aes(y=pilot_c2), colour="green") + geom_point(aes(y=pilot_c2), colour="green") +
  ylab("Percentage of Times Mind Changed") + xlab("Q Value") +
  labs(title="Percentage of Minds Changed vs Q value")
```



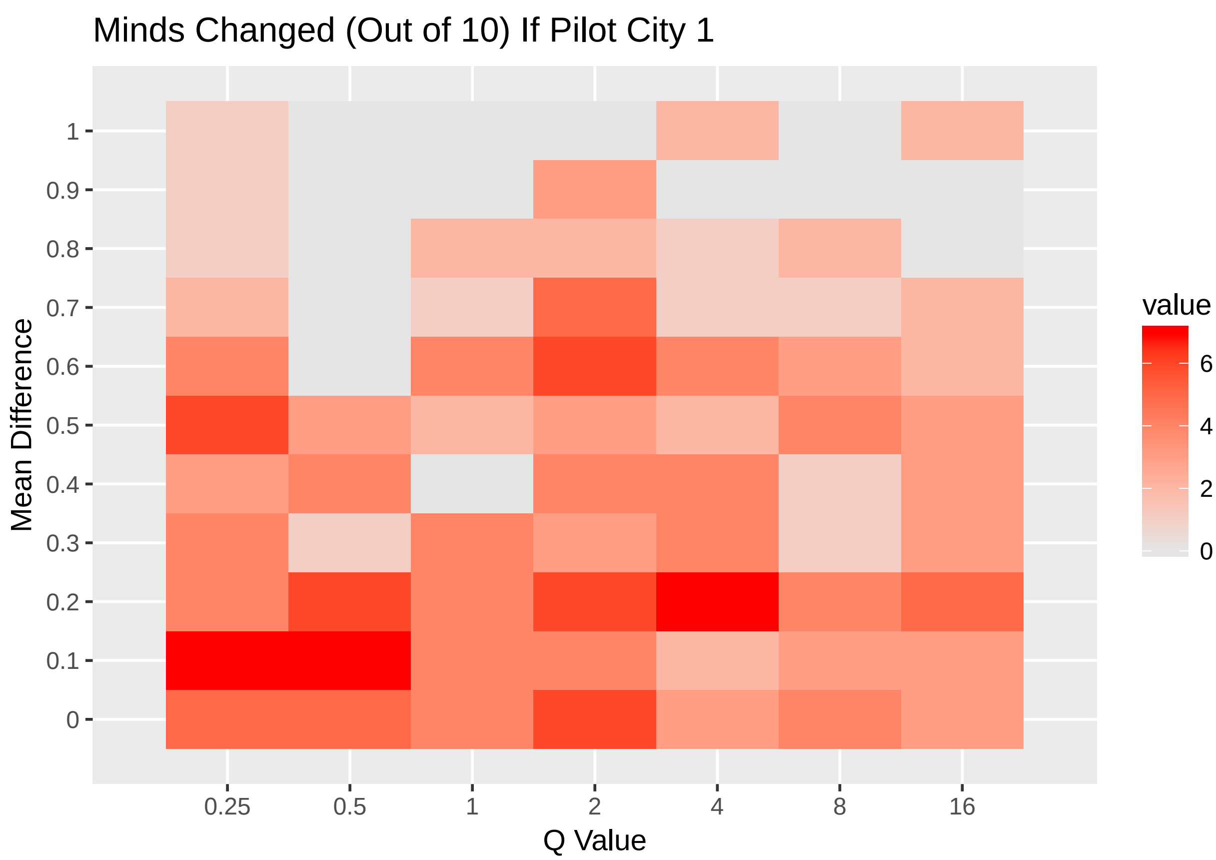
Here, we've again run a case where there are only two cities to chose from, we can run a pilot in only one city, and we can run the final program in only one city. City 1 has an initial prior $Y_1 \sim (1, 1)$ and city 2 has an initial prior $Y_2 \sim (0, 1)$ where x. What changes here is the variance of the pilot, where the pilot variance is Q times 'better' than the original study (A Q value of 5 corresponds with a pilot that has a fifth of the variance, a Q value of 1/2 means the pilot has twice the variance of the original study). Each simulation has been run 100 times, and the number of times changing the initial ranking (deciding to go for city 2 rather than city 1) has been plotted as the percentage of times we changed our minds (y). Piloting city 1 results plotted in blue, piloting city 2 results plotted in green.

```
Q_val <- c(0.25,0.5,1,2,4,8,16)
Q <- length(Q_val)
N <- 11
pilot_c1 <- matrix(nrow = N, ncol = Q)
pilot_c2 <- matrix(nrow = N, ncol = Q)
sample <- seq(0,1,length.out = N)
```

```
for (n in 1:N){
  for (q in 1:Q){
    data <- list(I=2,Y=c(sample[n],0),sigmaSq=c(1,1))
    results <- overall(data=data,num_pilots=1,num_final_cities=1,num_draws=10,Q=Q_val[q])
    pilot_c1[n,q] <- results$nmc[1]
    pilot_c2[n,q] <- results$nmc[2]
```

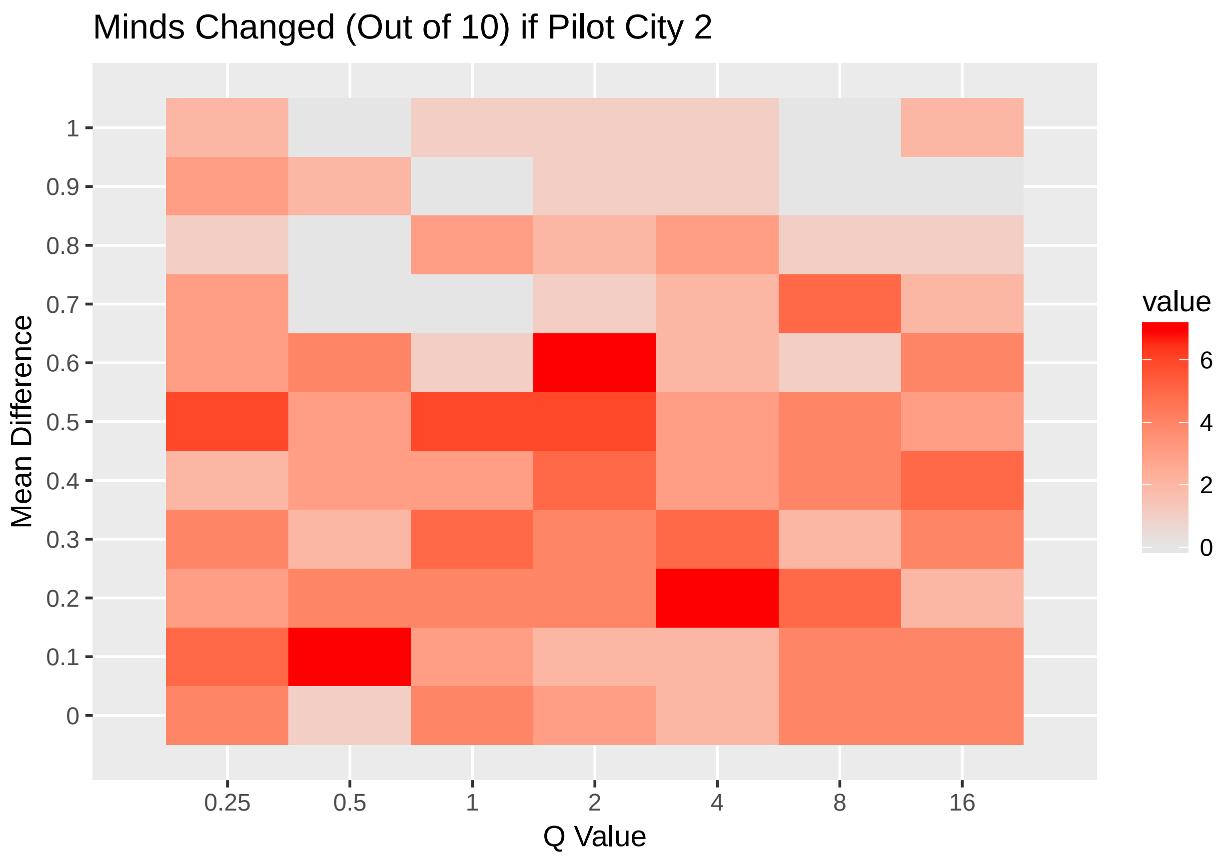
```
  }
}

df <- data.frame(pilot_c1, pilot_c2)
save(df, file="testing_Q_and_mean.Rda")
save(pilot_c1, file = "Q_and_mean_c1.Rdata")
save(pilot_c2, file = "Q_and_mean_c2.Rdata")
```

```
load("Q_and_mean_c1.Rdata")
longData<-melt(pilot_c1)

ggplot(longData, aes(x = Var2, y = Var1)) +
  geom_raster(aes(fill=value)) +
  scale_fill_gradient(low="grey90", high="red") +
  labs(x="Q Value", y="Mean Difference", title="Minds Changed (Out of 10) If Pilot City 1") +
  scale_x_discrete(name ="Q Value", limits=as.character(Q_val)) +
  scale_y_discrete(name ="Mean Difference", limits=as.character(sample))
```



Here we're trying to visualize how changing both the difference between the means of the two cities and the variance of the pilot studies (described by Q, with Q=3 implying the pilot variance is 1/3 of the original study variance). City 1 has prior mean x ranging fro 0 to 1 while city 2 has prior mean 0; both have prior variance 1. Here we can see that as the cities have closer means and the variance of the pilot study becomes larger, we're more likely to change our minds, as one would expect. Each square in this matrix represents the results of running the same simulation 10 times, counting the total number of times out of 10 that our mind is changed away from choosing city 1 as the final program destination, given that we choose to implement the pilot in city 1.

```r
load("Q_and_mean_c2.Rdata")
longData<-melt(pilot_c2)

ggplot(longData, aes(x = Var2, y = Var1)) +
  geom_raster(aes(fill=value)) +
  scale_fill_gradient(low="grey90", high="red") +
  labs(x="Q Value", y="Mean Difference", title="Minds Changed (Out of 10) if Pilot City 2") +
  scale_x_discrete(name ="Q Value", limits=as.character(Q_val)) +
  scale_y_discrete(name ="Mean Difference", limits=as.character(sample))
```



Here we're trying to visualize how changing both the difference between the means of the two cities and the variance of the pilot studies (described by Q, with Q=3 implying the pilot variance is 1/3 of the original study variance). City 1 has prior mean x ranging fro 0 to 1 while city 2 has prior mean 0; both have prior variance 1. Here we can see that as the cities have closer means and the variance of the pilot study becomes larger, we're more likely to change our minds, as one would expect. Each square in this matrix represents the results of running the same simulation 10 times, counting the total number of times out of 10 that our mind is changed away from choosing city 1 as the final program destination, given that we choose to implement the pilot in city 2.