auditStudy

Ryan Kennedy

10/22/2019

Summary

In compliance with the journal's requirement for log files for analysis, this R Markdown file contains a sample run of all the code that is in the auditStudy.R code that is included in the replication folder.

As noted several other places. The files provided omit IP Addresses, since some institutions consider these personally identifying data (PID). If your institution classifies IP addresses as non-PID these to be shared and you can contact the authors about sharing them. The R code in the replication file includes some commented code showing how we worked with this data and how you could do so in your own work.

Load Required Packages

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------- tidyverse 1.2.
1 --
```

```
## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------- tidyverse_conflicts
() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

Load the 37 Audited Studies

```
# Set Working Directory to the Replication Files folder
setwd("C:/Users/rkennedy/Dropbox/MTurk Cheating/Replication Files")

# Load the studies
files <- list.files("./Data/Audit Studies")

#########################
# Loop to load all of the data together
#########################
audit_data <- tibble() # Create an empty tibble
for(i in files) { # Loop over all the files in the folder
  temp_data <- read_csv(paste0("./Data/Audit Studies/",i)) # Load each individual dataset
  audit_data <- audit_data %>% # bind the data together
    bind_rows(temp_data)
}
```

```
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
```

```
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
```

```
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
```

```
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
```

```
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
```

```
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
```

```
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
```

```
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
```

```
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
```

```
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
## Parsed with column specification:
## cols(
##   StartDate = col_datetime(format = ""),
##   countryCode = col_character(),
##   countryName = col_character(),
##   asn = col_double(),
##   isp = col_character(),
##   block = col_double(),
##   hostname = col_character(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double()
## )
```

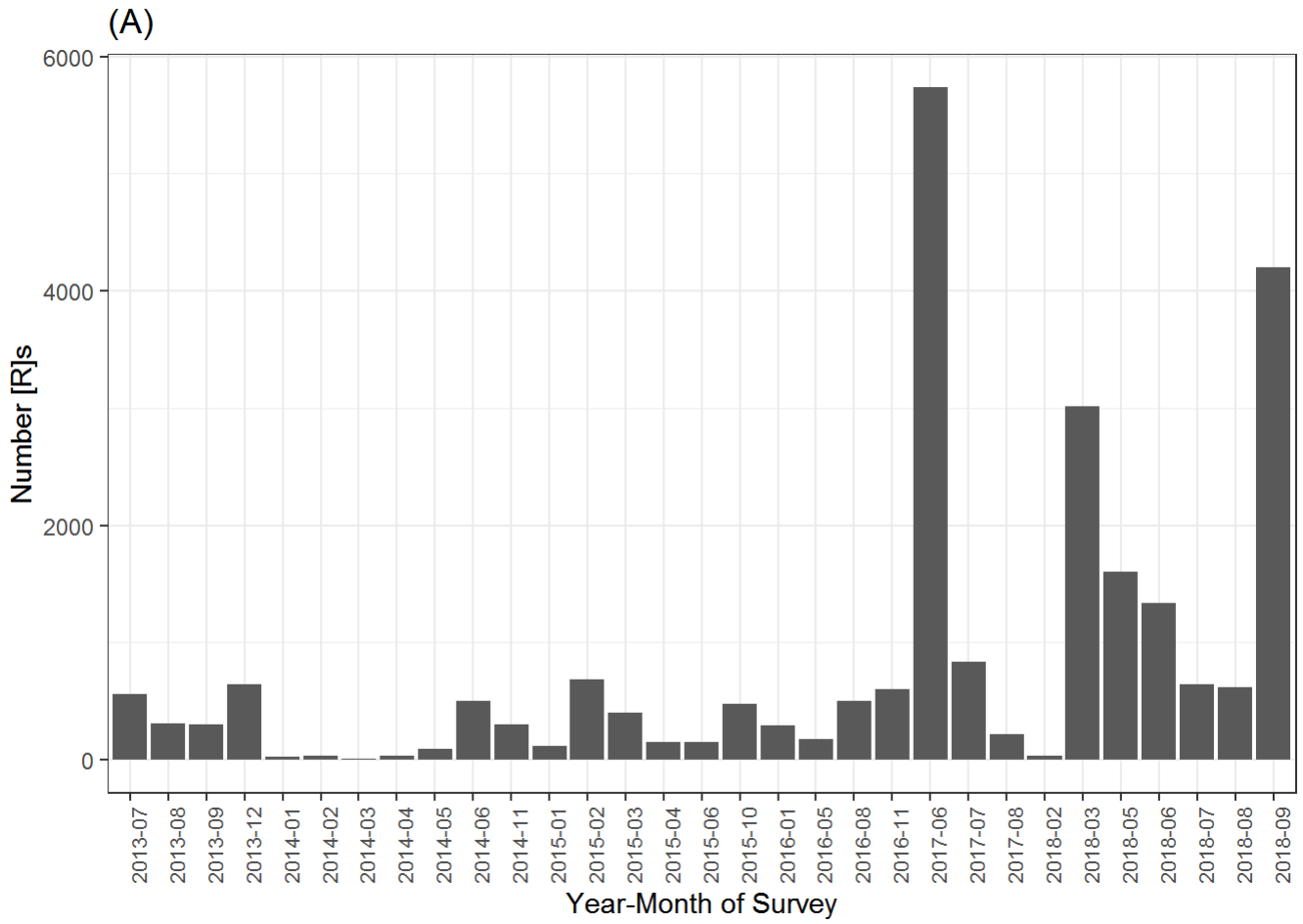
Create the Plots in Figure 1 and Combine Them

Start by setting up the data.

```
#########################
# Create plots to show the chart results
#########################

# Shorted name of Venezuela for chart
audit_data <- audit_data %>%
  mutate(countryName = recode(countryName, "Venezuela, Bolivarian Republic of" = "Venezuela"))

# Summarize the number of suspect responses by month
meanCheats <- audit_data %>%
  mutate(monthyear = format(StartDate, "%Y-%m"),
         VPS = ifelse(block == 1, 1, 0),
         nonUS = ifelse(countryCode != "US", 1, 0)) %>%
  group_by(monthyear) %>%
  summarise(pctVPS = mean(VPS, na.rm = T),
            pctNUS = mean(nonUS, na.rm = T),
            totResp = n())
```

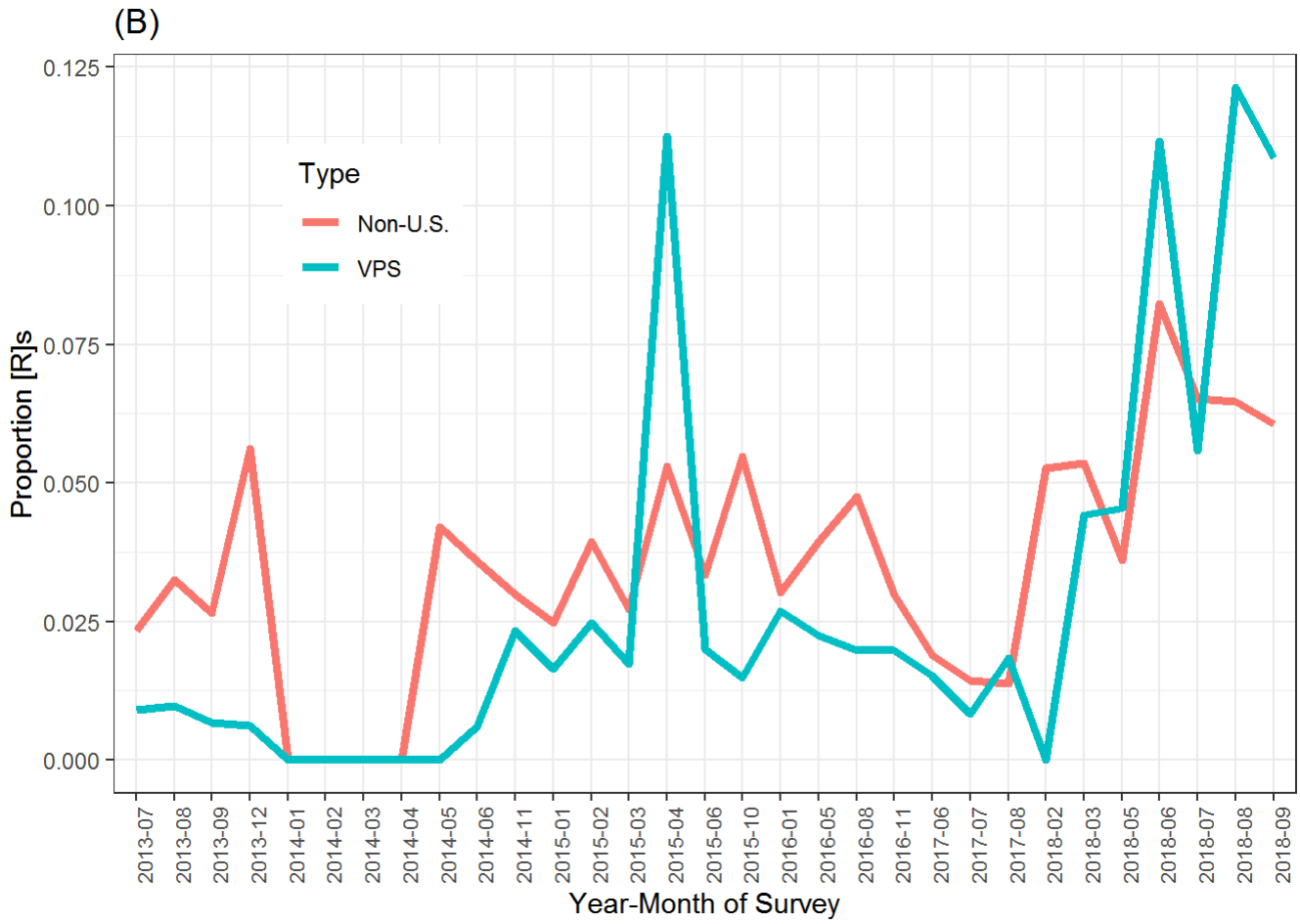
Plot A

```
# Plot A, the mean number of suspected respondents by month
p1 <- ggplot(meanCheats) +
  geom_bar(aes(x = monthyear, y = totResp), stat = "identity") +
  xlab("Year-Month of Survey") + ylab("Number [R]s") + theme_bw() +
  ggtitle("(A)") +
  theme(axis.text.x = element_text(angle = 90, size = 8))
p1
```



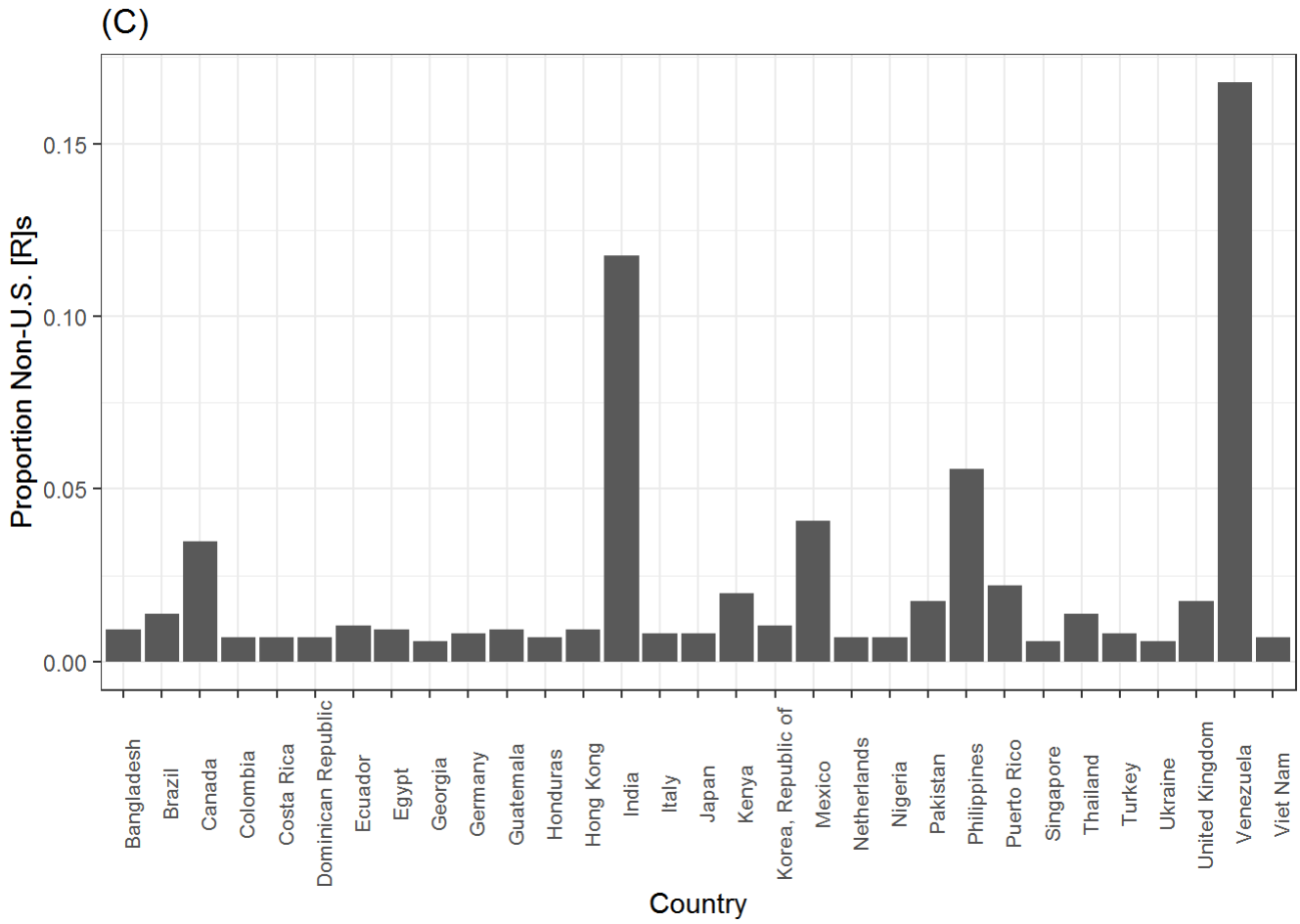
Plot B

```
# Plot B, the mean number, broken down by VPS and non-US over time
p2l <- meanCheats %>%
  gather(varName, proportion, pctVPS:pctNUS) %>%
  ggplot() +
  geom_line(aes(x = monthyear, y = proportion, group = varName, color = varName), size = 1.5) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 90, size = 8), legend.position = c(.22,.77)) +
  scale_color_discrete(name = "Type", labels = c("Non-U.S.", "VPS")) +
  ylab("Proportion [R]s") + xlab("Year-Month of Survey") +
  ggtitle("(B)")
p2l
```



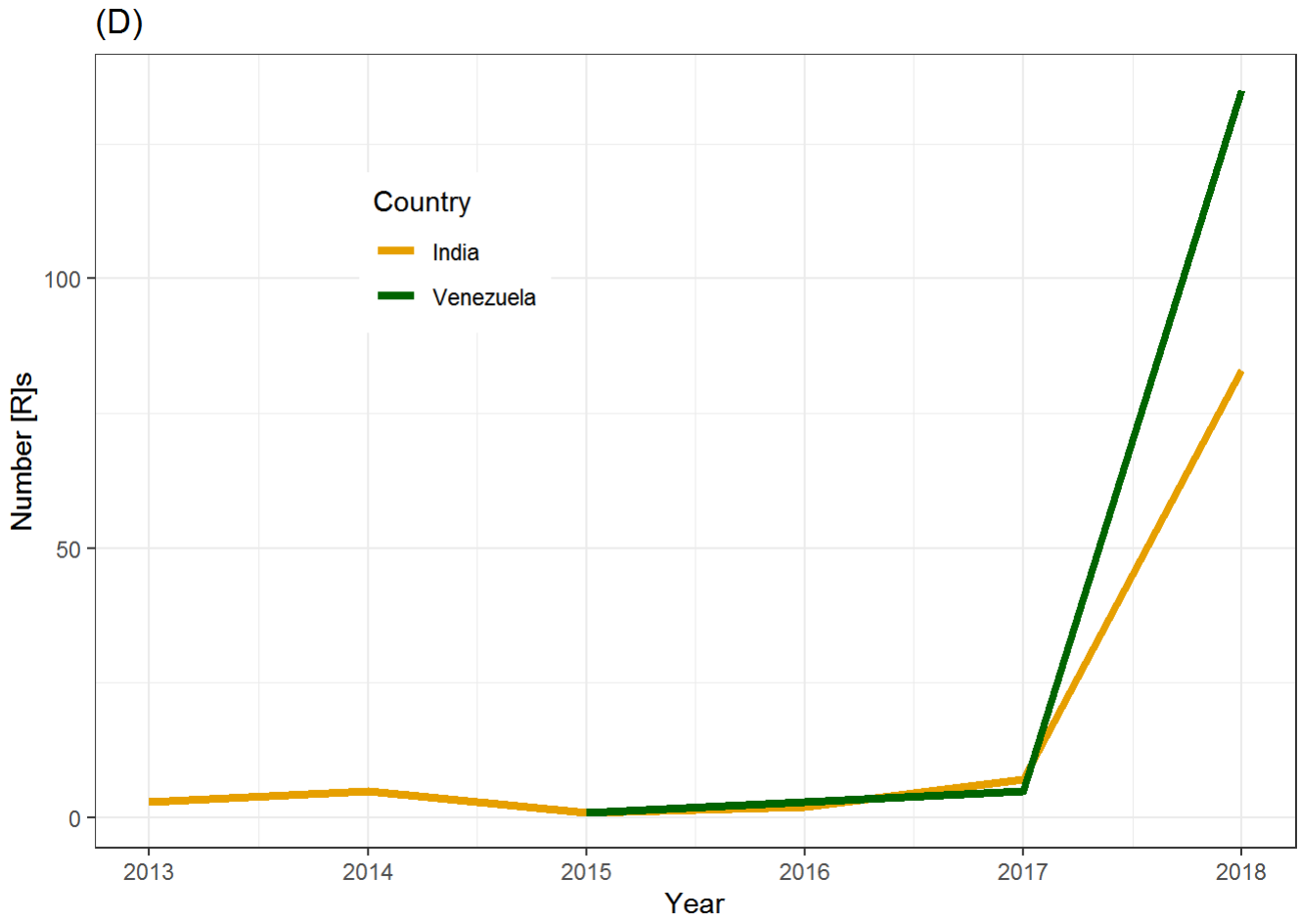
Plot C

```
# Plot C, the proportion of non-US respondents for countries with more than 4 respondents
p3 <- audit_data %>%
  filter(countryCode != "US" & block != 1) %>%
  count(countryName) %>%
  mutate(propCty = n/sum(n)) %>%
  filter(n > 4 & countryName != "Unknown") %>%
  ggplot() +
  geom_bar(aes(x = countryName, y = propCty), stat = "identity") + #coord_flip() +
  xlab("Country") + ylab("Proportion Non-U.S. [R]s") + theme_bw() +
  ggtitle("(C)") +
  theme(axis.text.x = element_text(angle = 90, size = 8))
p3
```



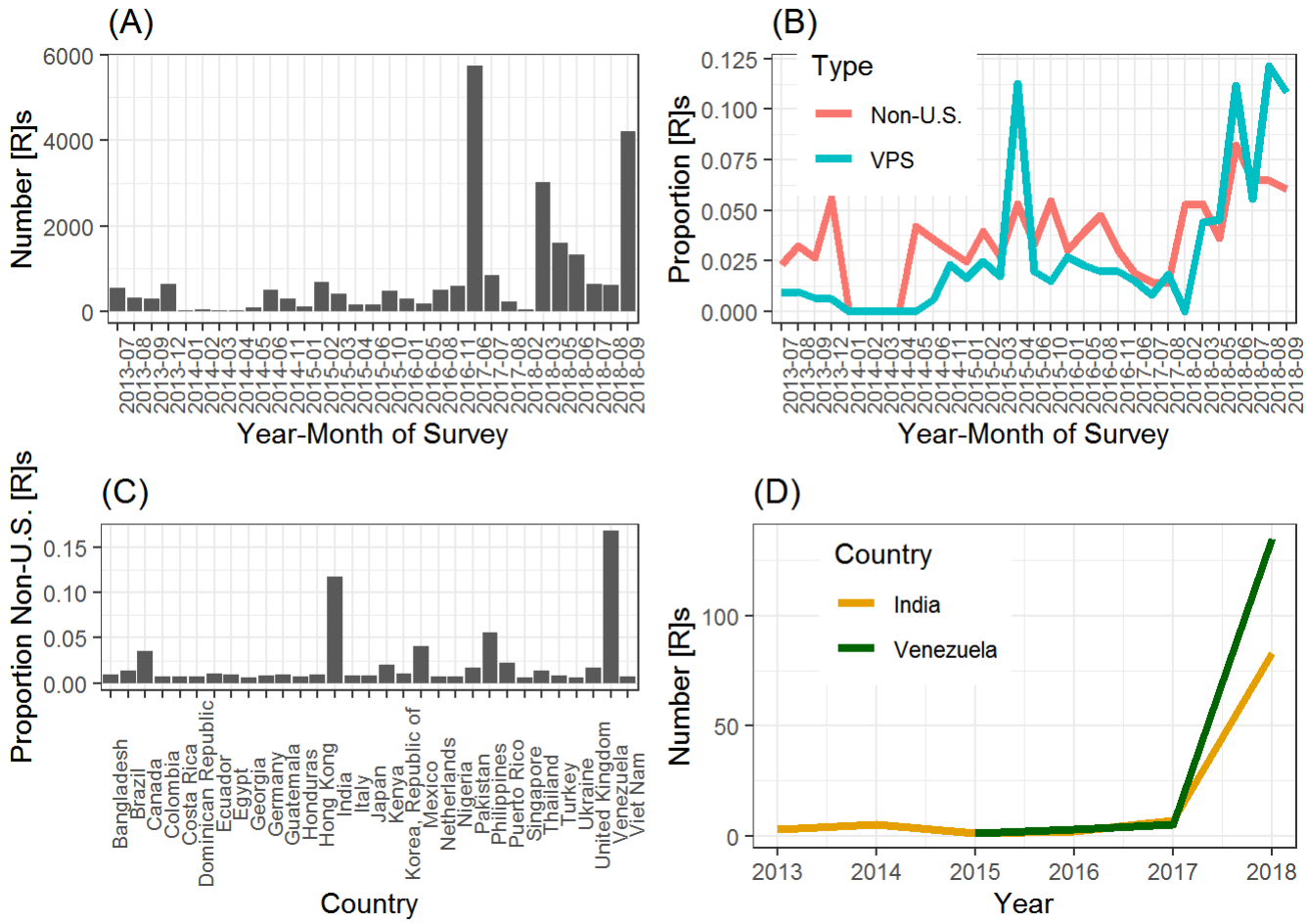
Plot D

```
# Create data with the number of resondents from Venezuela and India over time
timelineCheats <- audit_data %>%
  mutate(year = format(StartDate, "%Y")) %>%
  filter(countryName == "Venezuela" |
           countryName == "India") %>%
  group_by(year, countryName) %>%
  summarize(sumCheats = n())

# Plot D, number of respondents on US surveys that were actually from India or Venezuela
p4 <- timelineCheats %>%
  ggplot() +
  geom_line(aes(x = as.numeric(year), y = sumCheats, color = countryName), size = 1.5) +
  theme_bw() +
  scale_color_manual(name = "Country", labels = c("India","Venezuela"),
                     values = c("#E69F00", "darkgreen")) +
  xlab("Year") + ylab("Number [R]s") +
  ggtitle("(D)") +
  theme(legend.position = c(.3,.75))
p4
```



Combine and save the plots.

```
# Collect plots into grid for Figure 1 [Will need to adjust size to accomodate legends]
g <- grid.arrange(p1,p2l,p3,p4, ncol = 2)
```




```
# Write the plot to disk
ggsave("C:/Users/rkennedy/Dropbox/MTurk Cheating/Replication Files/Output Figures/timeseriesPlo
t.pdf", g, device = "pdf", height = 6, width = 8)
```

Comparison of IP Hub with Other Services

This is the companion code for Table 1. It starts from the compiled data, since data from the services was combined by IP address. Further information can be attained from the authors if needed.

Load the Data

```
##########################################
# Crosscheck IP Hub vs. Know Your IP
##########################################
# Load results from IP Hub, IP Void, and AbuseIPDB, and IP Hub data combined
# Note: since merging was done on IP Addresses, full merge data has been omitted
crosscheck <- read_csv("C:/Users/rkennedy/Dropbox/MTurk Cheating/Replication Files/Data/Comparis
on Data/crosscheck.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_character(),
##   StartDate = col_datetime(format = ""),
##   asn = col_double(),
##   block = col_double(),
##   vpsStd = col_double(),
##   vpsStrict = col_double(),
##   ipForeign = col_double(),
##   shldBlock = col_double(),
##   abuseipdb.found = col_double(),
##   iphubUS = col_double(),
##   iphubVPS = col_double(),
##   ipvoidUS = col_double(),
##   ipvoidVPS = col_double(),
##   abuseipdbUS = col_double(),
##   abuseipdbVPS = col_double()
## )
```

```
## See spec(...) for full column specifications.
```

Run the Comparisons

```
# Comparison of coding for US IPs
table(crosscheck$iphubUS, crosscheck$ipvoidUS)
```

```
##
##       0   1
##   0 534   4
##   1  10  28
```

```
table(crosscheck$iphubUS, crosscheck$abuseipdbUS)
```

```
##
##       0   1
##   0 534   4
##   1   3  35
```

```
# Comparison of coding for VPS use
table(crosscheck$iphubVPS, crosscheck$ipvoidVPS)
```

```
##
##       0   1
##   0 452  53
##   1  32  39
```

```
table(crosscheck$iphubVPS, crosscheck$abuseipdbVPS)
```

```
##
##       0   1
##   0 497   8
##   1   5  66
```

```
# Table 1 results for all exclusion criteria
table(crosscheck$iphubExclude, crosscheck$ipvoidExclude)
```

```
##
##              ipvoidBlock ipvoidSafe
##   iphubBlock          67         29
##   iphubSafe           37        443
```

```
table(crosscheck$iphubExclude, crosscheck$abuseipdbExclude)
```

```
##
##              abuseipdbBlock abuseipdbSafe
##   iphubBlock             91            5
##   iphubSafe              10          470
```

Check the Timing of Responses

To help determine whether the problematic responses were coming repeatedly from the same, or a small group, of individual users, we plotted the timing of responses in the Appendix. This code replicates that analysis.

```
##################################
# Check of timing of completion to see if this is an issue of a few users taking the survey many
times
# Figures in Section A7 of Appendix
##################################

# Load needed libraries
library(haven)
library(tidyverse)
library(lubridate)
```

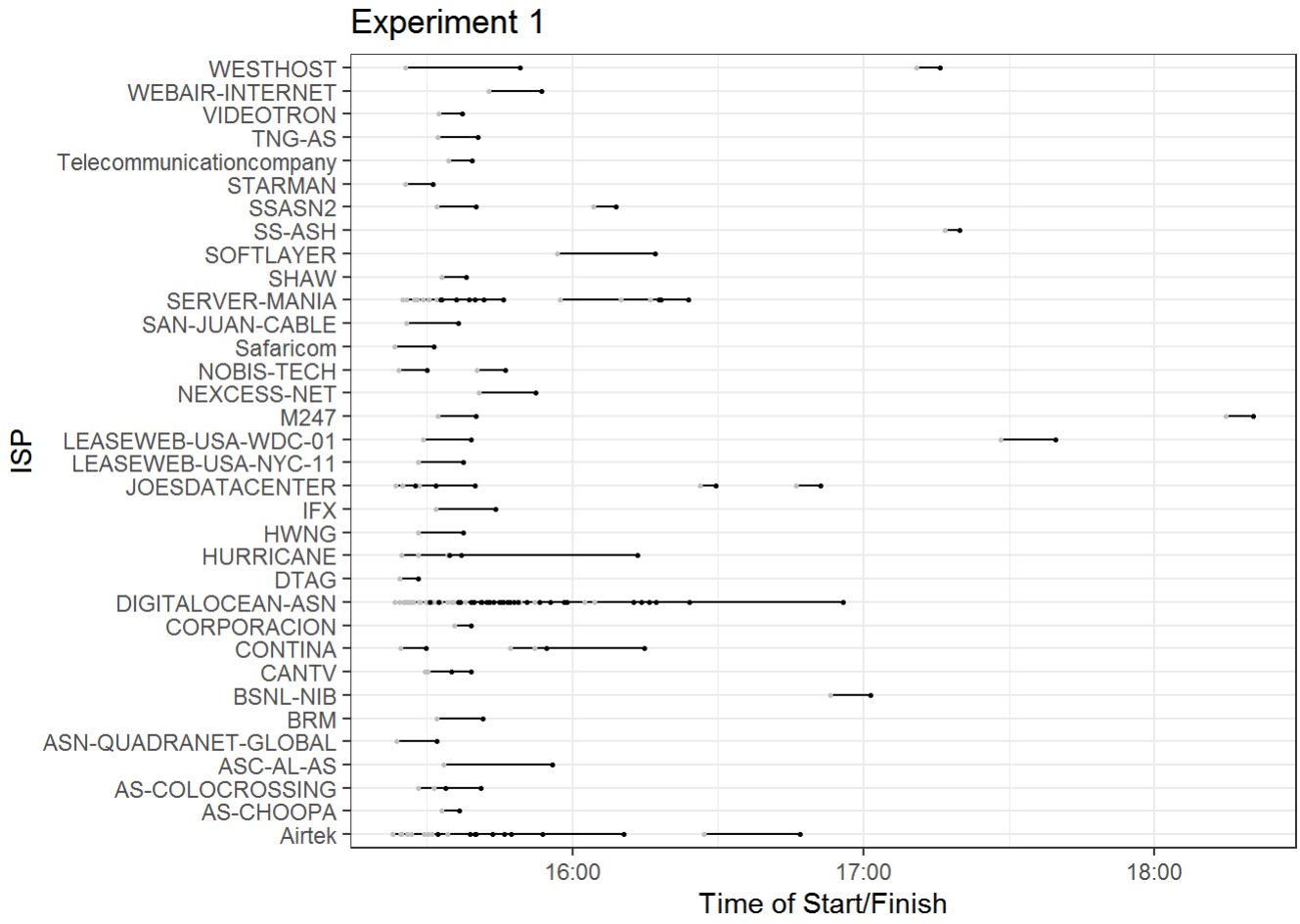
```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

```
library(ggalt)
```

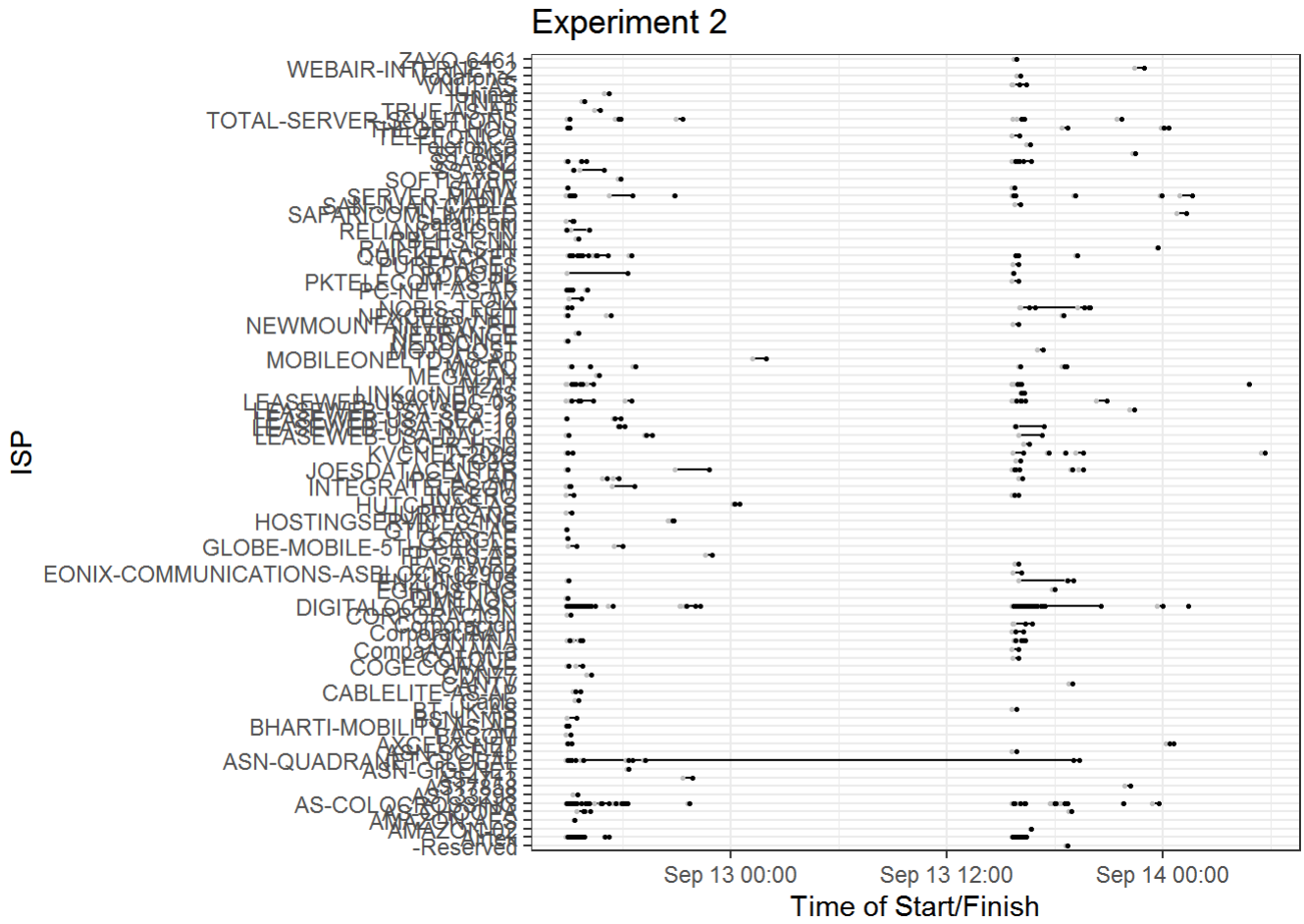
```
## Registered S3 methods overwritten by 'ggalt':
##   method                  from
##   grid.draw.absoluteGrob  ggplot2
##   grobHeight.absoluteGrob ggplot2
##   grobWidth.absoluteGrob  ggplot2
##   grobX.absoluteGrob      ggplot2
##   grobY.absoluteGrob      ggplot2
```

```
# Load data from teh two experiments
retro1 <- read_dta("C:/Users/rkennedy/Dropbox/MTurk Cheating/Replication Files/Data/retrostudy1
 data.dta")
retro2 <- read_dta("C:/Users/rkennedy/Dropbox/MTurk Cheating/Replication Files/Data/retrostudy2
 data.dta")

# Convert first experiment to only include those who finished the survey and were deemed suspect
by their IP address and format the dates for the plot
retro1 <- retro1 %>%
  filter(Finished == 1 & (block == 1 | countrycode != "US")) %>%
  mutate(start = ymd_hms(StartDate),
         end = ymd_hms(EndDate))

# Plot the start and finish times by ISP provider for suspect IP Addresses
tp1 <- ggplot(retro1) +
  geom_dumbbell(aes(x = start, xend = end, y = as.factor(isp)),
                colour_x = "grey", colour_xend = "black") +
  theme_bw() + labs(x = "Time of Start/Finish", y = "ISP", title = "Experiment 1")
tp1
```




```
# Write plot to disk
ggsave("C:/Users/rkennedy/Dropbox/MTurk Cheating/Replication Files/Output Figures/timingE1.pdf",
tp1, device = "pdf", height = 8, width = 6)
```

```
# Same data converstion for experiment 2
retro2 <- retro2 %>%
  filter(Finished == 1 & shldblock == 1) %>%
  mutate(start = ymd_hms(StartDate),
         end = ymd_hms(EndDate))

# Plot for experiment 2
tp2 <- ggplot(retro2) +
  geom_dumbbell(aes(x = start, xend = end, y = as.factor(isp)),
                colour_x = "grey", colour_xend = "black") +
  theme_bw() + labs(x = "Time of Start/Finish", y = "ISP", title = "Experiment 2")
tp2
```




```
# Write plot to disk (Note warnings are suppressed due to ASCII conversion warnings.)
suppressWarnings(ggsave("C:/Users/rkennedy/Dropbox/MTurk Cheating/Replication Files/Output Figur
es/timingE2.pdf", tp2, device = "pdf", height = 10, width = 8))
```