

Code for “To what extent did the introduction of TV advertisements in 1968 cause French newspapers to price discriminate?”

Mahmoud Elsheikh

December 22, 2020

```
# Code

main_dataset <- read_stata("/Users/mahmoudelsheikh/Desktop/STA304 Final Project/116438-V1/data/dta/main_dataset.d
ta")

maindata <- read_stata("/Users/mahmoudelsheikh/Desktop/STA304 Final Project/116438-V1/data/dta/main_dataset.dta")

# Dividing the subscription price by the unit price to obtain the price ratio
main_dataset$price_ratio <- main_dataset$ln_ps_cst / main_dataset$ln_po_cst

# Adding a share of unit buyers to the data set
main_dataset$unit_p <- 100 - main_dataset$q_s_s

main_dataset$year <- as.integer(main_dataset$year)

#Removing NA's from the data set
main_dataset <- main_dataset[!is.na(main_dataset$price_ratio),]
main_dataset <- main_dataset[!is.na(main_dataset$ra_cst),]
main_dataset <- main_dataset[!is.na(main_dataset$ra_s),]

# Creating summary statistics table

table1::label(main_dataset$po_cst) <- "Unit Price"
table1::label(main_dataset$ps_cst) <- "Subscription Price"
table1::label(main_dataset$price_ratio) <- "Price Ratio"
table1::label(main_dataset$ra_s) <- "Share of advertising revenues (%)"
table1::label(main_dataset$q_s_s) <- "Share of subscribers (%)"
table1::label(main_dataset$unit_p) <- "Share of unit buyers (%)"
table1::table1(po_cst + ps_cst + price_ratio + ra_s + q_s_s + ps_cst + unit_p, data = main_dataset)
```

	Overall (N=1012)
Unit Price	
Mean (SD)	3.25 (0.870)
Median [Min, Max]	3.27 [0.818, 9.35]
Subscription Price	
Mean (SD)	2.79 (0.725)
Median [Min, Max]	2.84 [0.682, 5.63]
Price Ratio	
Mean (SD)	0.843 (0.358)
Median [Min, Max]	0.872 [-9.22, 2.24]
Share of advertising revenues (%)	
Mean (SD)	46.3 (10.9)
Median [Min, Max]	45.9 [7.09, 81.0]
Share of subscribers (%)	
Mean (SD)	26.8 (22.2)
Median [Min, Max]	22.5 [0.699, 100]
Share of unit buyers (%)	
Mean (SD)	73.2 (22.2)
Median [Min, Max]	77.5 [-0.121, 99.3]

```
table6 <- table1(po_cst + ps_cst + price_ratio + ra_s + q_s_s + ps_cst + unit_p, data = main_dataset)

## Both Newspapers
both_news_ratio <- subset(main_dataset, main_dataset$year >= 1960, select=c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, po_cst, ps_cst, ra_s, q_s_s))
both_news_avg <- aggregate(both_news_ratio[, 1:8], list(both_news_ratio$year), mean)

## Local Newspapers

# Creating a data set combining all local newspapers data
local_news_ratio <- subset(main_dataset, main_dataset$year >= 1960 & main_dataset$local == 1, select=c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
local_news_avg <- aggregate(local_news_ratio[, 1:9], list(local_news_ratio$year), mean)

# Creating a data set combining all local newspapers data pre introduction of advertisements on TV (1966-1968)
local_news_ratio_pre_ads <- subset(main_dataset, main_dataset$year >= 1966 & main_dataset$year <= 1968 & main_data
set$local == 1, select=c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
local_news_pre_ads_avg <- aggregate(local_news_ratio_pre_ads[, 1:9], list(local_news_ratio_pre_ads$year), mean)

# Creating a data set combining all local newspapers data in the short-run introduction of advertisements on TV (1969-1971)
local_news_ratio_short_ads <- subset(main_dataset, main_dataset$year >= 1969 & main_dataset$year <= 1971 & mai
n_data$set$local == 1, select=c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
local_news_short_ads_avg <- aggregate(local_news_ratio_short_ads[, 1:9], list(local_news_ratio_short_ads$year), m
ean)

# Creating a data set combining all local newspapers data in the long-run introduction of advertisements on TV (1
972+)
local_news_ratio_long_ads <- subset(main_dataset, main_dataset$year >= 1972 & main_dataset$local == 1, select=c(r
total_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
local_news_long_ads_avg <- aggregate(local_news_ratio_long_ads[, 1:9], list(local_news_ratio_long_ads$year), mea
n)
```

```
## National newspapers

# Creating a data set combining all national newspapers data
national_news_ratio <- subset(main_dataset, main_dataset$year >= 1960 & main_dataset$local == 0, select=c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
national_news_avg <- aggregate(national_news_ratio[, 1:9], list(national_news_ratio$year), mean, drop = TRUE)

# Creating a data set combining all national newspapers data pre introduction of advertisements on TV (1966-1968)
national_news_ratio_pre_ads <- subset(main_dataset, main_dataset$year >= 1966 & main_dataset$year <= 1968 & main_data
set$local == 0, select=c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
national_news_pre_ads_avg <- aggregate(national_news_ratio_pre_ads[, 1:9], list(national_news_ratio_pre_ads$yea
r), mean, drop = TRUE)

# Creating a data set combining all national newspapers data in the short-run introduction of advertisements on T
V (1969-1971)
national_news_ratio_short_ads <- subset(main_dataset, main_dataset$year >= 1969 & main_dataset$year <= 1971 & mai
n_data$set$local == 0, select=c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
national_news_short_ads_avg <- aggregate(national_news_ratio_short_ads[, 1:9], list(national_news_ratio_short_ads
$year), mean, drop = TRUE)

# Creating a data set combining all national newspapers data in the long-run introduction of advertisements on TV
(1972+)
national_news_ratio_long_ads <- subset(main_dataset, main_dataset$year >= 1972 & main_dataset$local == 0, select=
c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
national_news_long_ads_avg <- aggregate(national_news_ratio_long_ads[, 1:9], list(national_news_ratio_long_ads$yea
r), mean, drop = TRUE)
```

```
# Effect of decrease in advertisement revenue on price ratio
combined_price_ratio <- data.frame(national_news_pre_ads_avg$price_ratio, local_news_pre_ads_avg$price_ratio, nat
ional_news_short_ads_avg)

combined_price_ratio1 <- data.frame(national_news_avg$price_ratio, local_news_avg$price_ratio, national_news_avg
$group.1)

# A least square regression model
lmodel_national <- lm(price_ratio ~ group.1, data = national_news_avg)
lmodel_local <- lm(price_ratio ~ group.1, data = local_news_avg)
```

```
# Combined newspaper linear regression model
lmodel_combined <- lm(formula = price_ratio ~ year + q_s_s + ra_s + ra_cst, data = both_news_ratio)

# National newspaper linear regression model post 1968
national_news_ratio_post1968 <- subset(main_dataset, main_dataset$year >= 1968 & main_dataset$local == 0, select=
c(rtotal_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
national_news_avg_post1968 <- aggregate(national_news_ratio_post1968[, 1:9], list(national_news_ratio_post1968$yea
r), mean, drop = TRUE)

local_news_ratio_post1968 <- subset(main_dataset, main_dataset$year >= 1968 & main_dataset$local == 1, select=c(r
total_cst, price_ratio, local, ln_po_cst, id_news, year, ra_cst, ra_s, q_s_s))
local_news_avg_post1968 <- aggregate(local_news_ratio_post1968[, 1:9], list(local_news_ratio_post1968$year), mea
n, drop = TRUE)

lmodel_national_post1968 <- lm(price_ratio ~ year + ra_cst + ra_s + q_s_s, data = national_news_ratio_post1968)
lmodel_local_post1968 <- lm(price_ratio ~ year + ra_cst + ra_s + q_s_s, data = local_news_ratio_post1968)
```

```
# Least square regression models for local and national newspapers in each period
lmodel_national_pre_ads <- lm(formula = price_ratio ~ year + ra_cst + ra_s + q_s_s, data = national_news_ratio_pr
e_ads)
lmodel_local_pre_ads <- lm(formula = price_ratio ~ year + ra_cst + ra_s + q_s_s, data = local_news_ratio_pre_ads)
lmodel_national_short <- lm(formula = price_ratio ~ year + ra_cst + ra_s + q_s_s, data = national_news_ratio_shor
t_ads)
lmodel_local_short <- lm(formula = price_ratio ~ year + ra_cst + ra_s + q_s_s, data = local_news_ratio_short_ads)
lmodel_national_long <- lm(formula = price_ratio ~ year + ra_cst + ra_s + q_s_s, data = national_news_ratio_long_
ads)
lmodel_local_long <- lm(formula = price_ratio ~ year + ra_cst + ra_s + q_s_s, data = local_news_ratio_long_ads)
```

```
# Table 1: National and Local newspaper models post 1968
table1 <- tab_model(lmodel_national_post1968, lmodel_local_post1968,
dv.labels = c("National Newspapers Post 1968", "Local Newspapers Post 1968"),
pred.labels = c("Intercept", "Year", "Advertising Revenues", "Share Of Advertising Revenues(%)", "Share of subscrib
ers(%)"
))

# Table 2: National and Local newspaper model pre TV advertisements (1966-1968)
table2 <- tab_model(lmodel_national_pre_ads, lmodel_local_pre_ads,
dv.labels = c("National Pre TV Ads", "Local Pre TV Ads"),
pred.labels = c("Intercept", "Year", "Advertising Revenues", "Share Of Advertising Revenues
(%)", "Share of subscribers(%)"
))
```

	National Pre TV Ads			Local Pre TV Ads		
Predictors	Estimates	CI	p	Estimates	CI	p
Intercept	-10.39	-68.43 – 47.66	0.717	24.57	-245.69 – 294.83	0.858
Year	0.01	-0.02 – 0.04	0.696	-0.01	-0.15 – 0.13	0.868
Advertising Revenues	-0.00	-0.00 – 0.00	0.666	0.00	-0.00 – 0.00	0.332
Share Of Advertising Revenues(%)	0.00	-0.00 – 0.00	0.308	-0.02	-0.03 – 0.01	0.002
Share of subscribers(%)	0.00	-0.00 – 0.00	0.232	-0.00	-0.01 – 0.00	0.077
Observations	33			195		
R ² / R ² adjusted	0.069 / -0.064			0.076 / 0.056		

Table 3: National and Local newspaper model in the short-run period of TV advertisements introduction (1969-1971)

```
table3 <- tab_model(lmodel_national_short, lmodel_local_short,
dv.labels = c("National Short-run TV Ads", "Local Short-run TV Ads"),
pred.labels = c("Intercept", "Year", "Advertising Revenues", "Share Of Advertising Revenues(%)",
"Share of subscribers(%)"
))
```

	National Short-run TV Ads			Local Short-run TV Ads		
Predictors	Estimates	CI	p	Estimates	CI	p
Intercept	29.92	5.42 – 54.43	0.019	-36.79	-119.83 – 46.25	0.383
Year	-0.01	-0.03 – 0.00	0.022	0.02	-0.02 – 0.06	0.366
Advertising Revenues	0.00	-0.00 – 0.00	0.914	0.00	-0.00 – 0.00	0.068
Share Of Advertising Revenues(%)	0.00	-0.00 – 0.00	0.064	-0.01	-0.02 – 0.01	<0.001
Share of subscribers(%)	0.00	0.00 – 0.00	<0.001	-0.00	-0.00 – 0.00	0.030
Observations	30			193		
R ² / R ² adjusted	0.598 / 0.533			0.139 / 0.120		

Table 4: National and Local newspaper model in the long-run period of TV advertisements introduction (1972-1974)

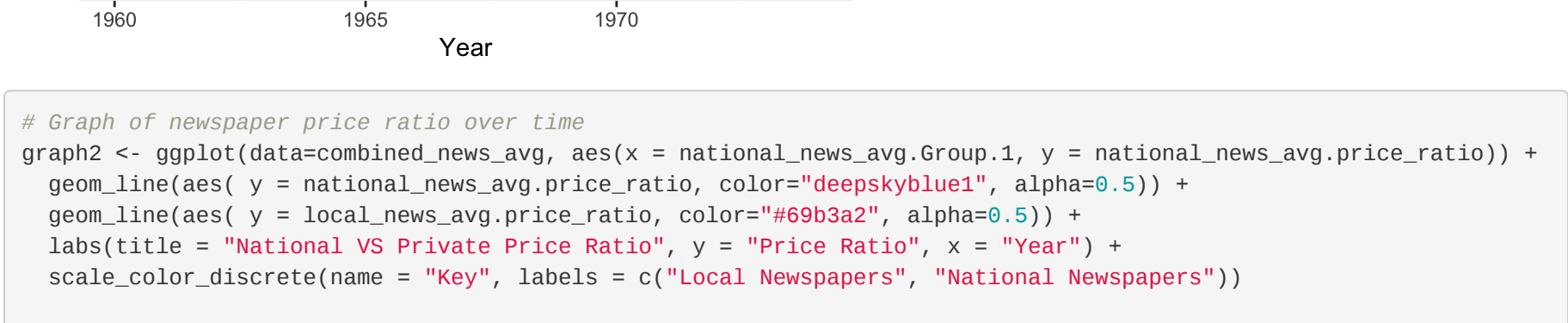
```
table4 <- tab_model(lmodel_national_long, lmodel_local_long,
dv.labels = c("National Long-run TV Ads", "Local Long-run TV Ads"),
pred.labels = c("Intercept", "Year", "Advertising Revenues", "Share Of Advertising Revenues(%)",
"Share of subscribers(%)"
))
```

	National Long-run TV Ads			Local Long-run TV Ads		
Predictors	Estimates	CI	p	Estimates	CI	p
Intercept	4.81	-41.33 – 50.95	0.830	-4.72	-38.73 – 29.29	0.784
Year	-0.00	-0.03 – 0.02	0.856	0.00	-0.01 – 0.02	0.738
Advertising Revenues	0.00	-0.00 – 0.00	0.066	0.00	0.00 – 0.00	0.001
Share Of Advertising Revenues(%)	0.00	-0.00 – 0.00	0.901	-0.00	-0.01 – 0.00	<0.001
Share of subscribers(%)	0.00	-0.00 – 0.00	0.603	-0.00	-0.00 – 0.00	0.205
Observations	25			184		
R ² / R ² adjusted	0.228 / 0.073			0.124 / 0.104		

```
combined_news_avg <- data.frame(national_news_avg$group.1, national_news_avg$price_ratio, local_news_avg$price_ra
tio, national_news_avg$ra_cst, local_news_avg$ra_s)

# Graph of share of advertising revenues for local and national newspapers
graph1 <- ggplot(data=combined_news_avg, aes(x = national_news_avg.group.1, y = national_news_avg.ra_s)) +
geom_line(aes(y = national_news_avg.ra_s, color="deepskyblue1", alpha=0.5)) +
geom_line(aes(y = local_news_avg.ra_s, color="red3", alpha=0.5)) +
labs(title = "National VS Private Newspaper Advertisement Revenue", y = "Share of Advertising Revenue (%)", x =
"Year") +
scale_color_discrete(name = "Key", labels = c("Local Newspapers", "National Newspapers"))

graph1
```



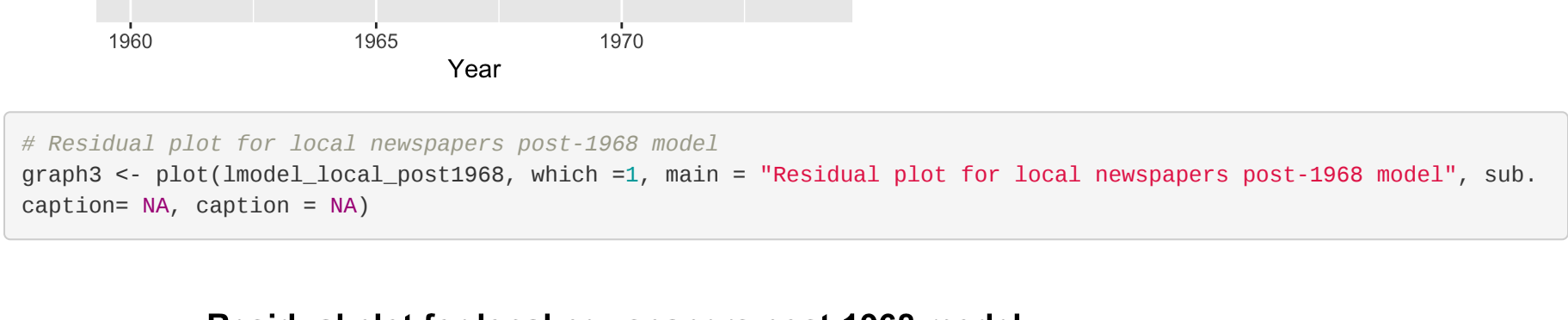
```
# Graph of newspaper price ratio over time
graph2 <- ggplot(data=combined_news_avg, aes(x = national_news_avg.group.1, y = national_news_avg.price_ratio)) +
geom_line(aes(y = national_news_avg.price_ratio, color="deepskyblue1", alpha=0.5)) +
geom_line(aes(y = local_news_avg.price_ratio, color="red3", alpha=0.5)) +
labs(title = "National VS Private Price Ratio", y = "Price Ratio", x = "Year") +
scale_color_discrete(name = "Key", labels = c("Local Newspapers", "National Newspapers"))

graph2
```



Residual plot for local newspapers post-1968 model

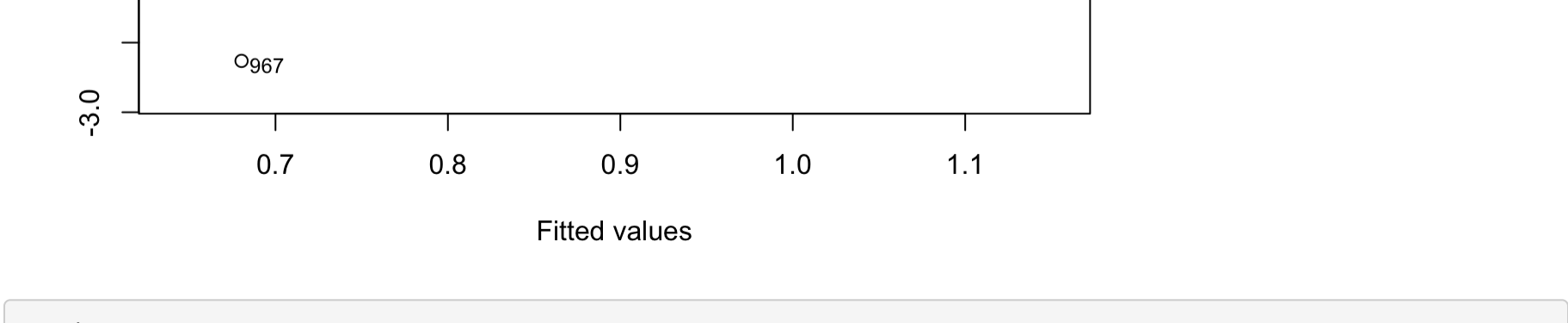
```
graph3 <- plot(model_national_post1968, which=1, main = "Residual plot for local newspapers post-1968 model", sub.
caption= NA, caption = NA)
```



```
graph3
```

```
## NULL
```

```
# Residual plot for national newspapers post-1968 model
graph4 <- plot(model_national_post1968, which=1, main = "Residual plot for national newspapers post-1968 model", sub.
caption= NA, caption = NA)
```

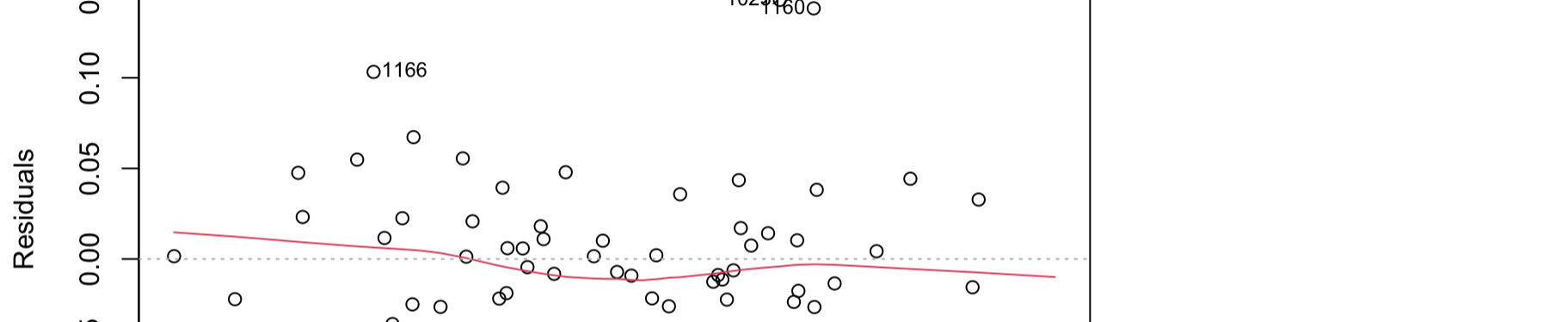


```
graph4
```

```
## NULL
```

```
# Plot of advertising revenues share for local and national newspapers
graph5 <- (main_dataset %>%
mutate(type = if_else(local == 1, "Local", "National"))) %>%
ggplot(aes(x = year, y = ra_s)) +
geom_point(alpha = 0.5) +
scale_y_continuous() +
labs(title = "Plot of advertising revenues share for local and national newspapers",
x = "Year",
y = "Share of Advertising Revenues (%)") +
facet_wrap(vars(type),
nrow = 2) +
theme_minimal() +
geom_vline(xintercept = 1968.5, linetype = "dashed")

graph5
```



```
# Graph of price ratio vs advertising revenue
graph6 <- ggplot(model_combined, aes(y = price_ratio, x = 100-ra_s)) +
geom_point() +
geom_smooth(method="lm") +
ylim(0.5, 1.2) +
labs(title = "Price ratio VS share of advertising revenue", y = "Price Ratio", x = "Share of Advertising Revenue
(%)")

graph6
```

