

Diseño y Evaluación de RCTs

Isidoro Garcia

2021

```
library(tidyverse)
library(data.table)
library(broom)
library(knitr)
library(lubridate)
library(RCT)
```

Contexto

Rappi te contrata para hacer una intervención que active a sus usuarios en la plataforma. La empresa quiere evaluar si es necesario darles cash los usuarios (y cuánto) para reactivar a los usuarios o si es suficiente con una campaña informativa.

Para ello, te decides a realizar un experimento factorial donde evalúas:

- El impacto de mandar un mensaje informativo donde muestres las nuevas tiendas afiliadas a Rappi, y
- El impacto de dar 100 ó 200 pesos en cupones

Finalmente, les gustaría entender cómo interactuar el mostrar las nuevas tiendas aunado con dar dinero en cupones.

A la empresa le gustaría entender el impacto de la intervención sobre:

- Las compras
- La tasa de usuarios activos en la app (transaccionar: abrir app o hacer compras)

Datos

Los datos para asignar los pueden encontrar en `universo.Rdata`.

Cargemos los datos

```
load('Bases input/universo.Rdata')
```

1. Cuántos grupos de tratamiento debe de haber? Elabora sobre que intervención va a recibir cada uno

- Control: Ninguna intervención
- Trat 1: Sólo mensaje
- Treat 2: Sólo cupón \$100
- Treat 3: Sólo cupón de \$200
- Treat 4: Sólo cupón de \$300
- Treat 5: Mensaje + \$100
- Treat 6: Mensaje + \$200
- Treat 7: Mensaje + \$300

2 (2pts). Como pueden notar, tenemos 2 poblaciones: Usuarios inactivos y usuarios que nunca estuvieron activos. Para ellos, las tasas de transaccionalidad son hasta ahora 7.94% y 0%. Utiliza esta información para hacer pruebas de poder: Dada esta tasa y población, cuál es el efecto mínimo detectable sobre la tasa de transaccionalidad como función de cuantas observaciones asignamos al grupo control? Interpreta. (Tip: asegurate de dejar claros los grupos comparados en esta prueba)

Vemos que con una grupo de control de 4 por ciento o más ya podemos detectar diferencias de menos de 1 punto porcentual entre el grupo de control y cada tratamiento. Esto nos da un termómetro de que tanto podemos hacer chico el grupo control (por razones de negocio) sin sacrificar tanto el poder estadístico.

```
# La tasa ponderada
grupos<-
  universo %>%
  group_by(population) %>%
  tally()

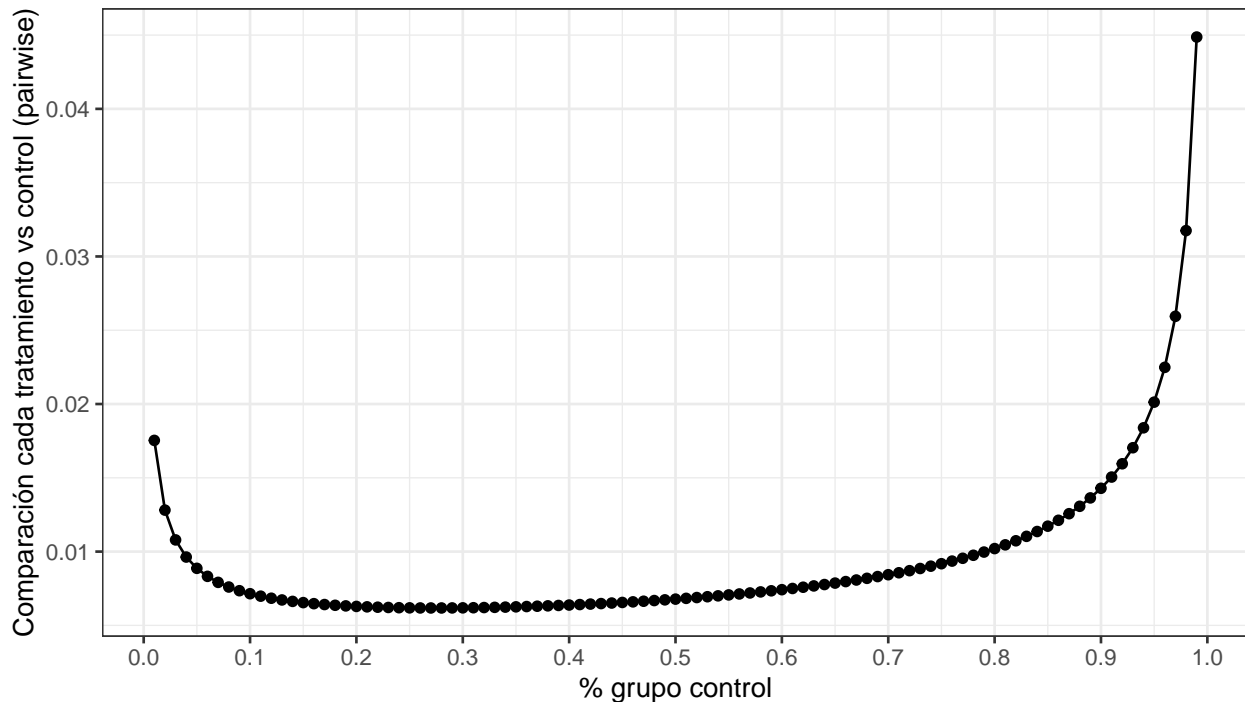
grupos$tasa<-c(0.0794,0)

prior <- weighted.mean(grupos$tasa, grupos$n)

# Prueba de poder sobre tasa de transacción
power_test<-
  tau_min_probability(prior = mean(grupos$tasa),
                      N = nrow(universo),
                      share_control = seq(0.01,0.99, 0.01),
                      n_groups = 8)

ggplot(power_test, aes(share_control, tau_min_each_treat))+
  geom_point()+
  geom_line()+
```

```
theme_bw()+labs(x = '% grupo control', y = 'Comparación cada tratamiento vs control (pairwise)')+
scale_x_continuous(breaks = seq(0,1,0.1))
```

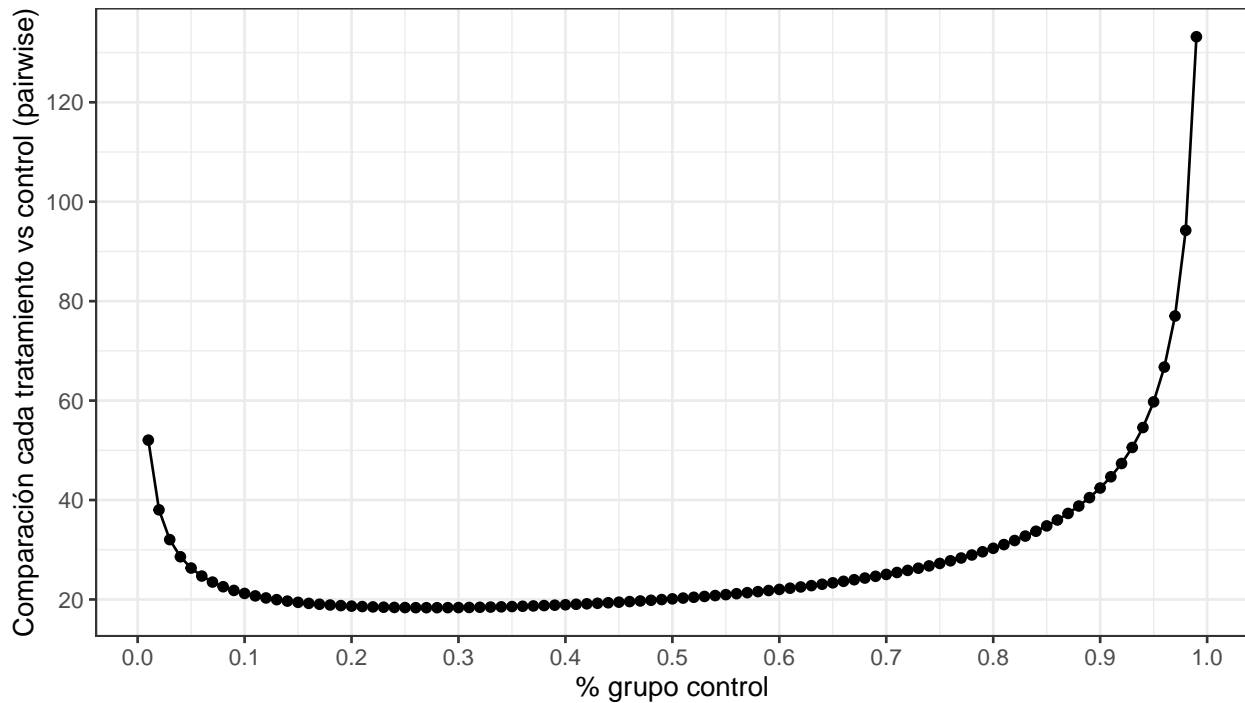


3. Repite el mismo ejercicio pero ahora para usando las compras totales como variable objetivo. Elige un share de control con base en tu respuesta de esta y la anterior pregunta

Si ahora hacemos las pruebas de poder sobre compras, necesitamos un control de 12 por ciento para detectar una diferencia de 20 pesos entre los grupos. Esto, dada la distribución original de las compras (mediana = 0, media = 54) parece ser la mejor idea.

```
power_test_compras<-
  tau_min(outcome_var = universo$total_purchases,
          N = nrow(universo),
          share_control = seq(0.01,0.99,0.01),
          n_groups = 8)

ggplot(power_test_compras, aes(share_control, tau_min_each_treat))+
  geom_point()+
  geom_line()+
  theme_bw()+labs(x = '% grupo control', y = 'Comparación cada tratamiento vs control (pairwise)')+
  scale_x_continuous(breaks = seq(0,1,0.1))+
  scale_y_continuous(breaks = seq(20, 150, 20))
```



4 (2ptos) Qué variables crees que puedan estar más correlacionadas con el impacto? Justifica tu respuesta y elige un set

Listemos las variables primero:

- Device value: El valor del celular puede ser un proxy de ingreso del usuario. Esto puede afectar en que tan relevante son los incentivos monetarios para el usuario y por ende con τ
- Población: Sin duda la actividad pasada es un gran indicador de la actividad futura.
- Teléfono verificado: El teléfono puede indicar fricciones. Si el usuario debe verificar el teléfono antes de comprar, en verdad le estamos pidiendo 2 cosas en lugar de 1.
- Genero: La literatura previa muestra que las mujeres reaccionan mejor a regalos que los hombres
- Edad: La edad puede influir mucho en qué tan tech savvy son los usuarios y por ende con su reacción a los tratamientos.

5 (2ptos) Realiza la asignación aleatoria. Muestra la distribución de los grupos por estrato, los misfits. Sin mirar el balance, lograron una asignación aleatoria exitosa? Justifica tu respuesta

```
# Estratifiquemos device_value, phone_verified, population
asignacion<-treatment_assign(data = universo,
                             share_control = 1/8,
                             n_t = 7,
                             strata_varlist = vars(device_value, phone_verified, population),
                             missfits = 'global', seed = 1900, key = "user_id")
```

```
Warning: Unknown or uninitialised column: `treat`.
Unknown or uninitialised column: `treat`.
```

```
kable(asignacion$summary_strata)
```

strata	device_value	phone_verification	population	n_stratan	missfits	0	1	2	3	4	5	6	7
1	[\$0-\$10,000]	0	Inactive	3484	4	436	435	436	436	435	436	435	435
2	[\$0-\$10,000]	0	Never Active	21408	0	2676	2676	2676	2676	2676	2676	2676	2676
3	[\$0-\$10,000]	1	Inactive	6155	3	769	769	769	769	769	771	769	770
4	[\$0-\$10,000]	1	Never Active	18632	0	2329	2329	2329	2329	2329	2329	2329	2329
5	[\$10,001+]	0	Inactive	5	5	1	NA	1	1	NA	NA	2	NA
6	[\$10,001+]	0	Never Active	35	3	4	4	5	4	6	4	4	4
7	[\$10,001+]	1	Inactive	9	1	1	1	1	1	1	1	1	2
8	[\$10,001+]	1	Never Active	27	3	3	4	4	3	3	3	4	3
9	No info	0	Inactive	2013	5	252	253	251	251	251	252	251	252
10	No info	0	Never Active	33523	3	4190	4190	4190	4192	4190	4190	4190	4191
11	No info	1	Inactive	2418	2	302	302	302	302	304	302	302	302
12	No info	1	Never Active	16482	2	2060	2061	2060	2060	2060	2060	2061	2060

```
universo<-
  left_join(universo, asignacion$data)

universo %>%
  group_by(treat) %>%
  tally() %>%
  kable()
```

treat	n
0	13023
1	13024
2	13024
3	13024
4	13024
5	13024
6	13024
7	13024

6. Qué elección tomaron sobre como manejar los misfits? Elaboren sus razones

Elegí `global` dado que en los casos de negocio, típicamente es muy importante cuidar que las proporciones de los grupos se cumplan. De otra manera los presupuestos pueden afectarse.

7. Realiza las pruebas de balance t sobre todas las variables (Tip: transforma las categóricas en dummies). Parece haber balance?

Si. Todas las variables para todos los grupos tienen p-values mayores a 10 por ciento. Por lo tanto, parece haber balance entre grupos (excludability)

```
library(fastDummies)

universo<-
  universo %>%
  dummy_cols()

balance_t<-balance_table(data = universo %>% select(-user_id) %>% select_if(is.numeric),
  treatment = "treat")

kable(balance_t %>% select(contains("Media")), digits = 2)
```

Media_control1	Media_trat1	Media_trat2	Media_trat3	Media_trat4	Media_trat5	Media_trat6	Media_trat7
29.12	28.98	29.07	29.11	29.04	29.31	29.05	29.16
0.48	0.48	0.48	0.48	0.48	0.48	0.48	0.48
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52
0.26	0.24	0.25	0.25	0.25	0.26	0.25	0.25
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4.81	4.86	4.87	4.83	4.86	4.79	4.88	4.83
0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
6.89	6.89	6.89	6.89	6.89	6.89	6.89	6.89
55.49	51.95	58.27	50.41	53.74	61.68	49.16	56.76

```
kable(balance_t %>% select(contains("p")), digits = 2)
```

p_value1	p_value2	p_value3	p_value4	p_value5	p_value6	p_value7
0.31	0.74	0.97	0.56	0.16	0.61	0.77
0.99	1.00	1.00	0.99	0.98	0.99	1.00
1.00	0.65	1.00	0.82	0.81	0.65	1.00
0.99	0.98	1.00	1.00	0.99	0.99	1.00
0.02	0.05	0.73	0.31	0.94	0.22	0.33
0.71	0.71	0.71	0.71	0.71	0.71	0.71
0.35	0.28	0.75	0.35	0.70	0.17	0.70
0.98	0.99	1.00	0.98	0.98	0.98	0.98
0.98	0.98	0.98	0.98	0.99	0.98	1.00
0.98	0.98	0.98	0.98	0.99	0.98	1.00
0.98	1.00	0.99	0.99	0.99	0.99	0.99
0.56	0.67	0.40	0.81	0.36	0.28	0.88

8. Repite el ejercicio pero ahora con pruebas de balance conjuntas. Muestra los resultados (incluyendo el estadístico de prueba, grados de libertad y p values) Interpreta

Al correr una regresión de cada tratamiento vs el control como variable endógena vs todas las variables explicativas. Obtenemos las pruebas conjuntas que indican si las características de los usuarios seleccionan (afectan la probabilidad) de pertenecer a cada grupo de tratamiento. Podemos ver que la regresión arroja $p > 0.1$ para todos los casos. Es decir, las características no sirven para determinar a qué grupo pertenece el usuario (o la regla es exógena).

```
balance_f<-balance_regression(data = universo %>% select(-user_id) %>% select_if(is.numeric),
                             treatment = "treat")

kable(balance_f$F_test)
```

estadístico	Msj1	Msj2	Msj3	Msj4	Msj5	Msj6	Msj7
F-statistic	7.685605e-01	6.062681e-01	1.195191e-01	2.444420e-01	3.539717e-01	5.086002e-01	1.637283e-01
k	9.000000e+00	9.000000e+00	9.000000e+00	9.000000e+00	9.000000e+00	9.000000e+00	9.000000e+00
n-k-1	2.603700e+04	2.603700e+04	2.603700e+04	2.603700e+04	2.603700e+04	2.603700e+04	2.603700e+04
F_critical	3.694309e-01	3.694309e-01	3.694309e-01	3.694309e-01	3.694309e-01	3.694309e-01	3.694309e-01
p_value	6.457549e-01	7.928312e-01	9.992421e-01	9.878919e-01	9.564626e-01	8.694666e-01	9.973371e-01
R	2.656000e-04	2.095000e-04	4.130000e-05	8.450000e-05	1.223000e-04	1.758000e-04	5.660000e-05
cuadrada							

9. Elabora porqué parecen cumplirse los 3 supuestos de la asignación

Si. Todas las variables para todos los grupos tienen p-values mayores a 10 por ciento. Por lo tanto, parece haber balance entre grupos. Esto nos da señales de que la asignación de tratamiento es 1) Probabilística (Overlap) y 2) Excludable. Más aún, dado que es un experimento digital, parece que podemos tener SUTVA, ya que los usuarios no deberían de contaminarse.

10. Elabora un pitch de negocio sobre los beneficios que este experimento podría dejar a Rappi.

Evaluación

Pasemos a la evaluación de tu intervención. En este ejercicio, Rappi diseñó un nuevo experimento con tus enseñanzas algo distinto al tuyo.

Este consistió en 6 grupos de tratamiento y un control:

- T1: Dar 100 pesos en cupones (con mensaje)
- T2: Dar 200 pesos en cupones (con mensaje)
- T3: Dar un descuento de 20% en la siguiente compra
- T4: Dar un descuento de 25% en la siguiente compra
- T5: Ofrecer 2% de descuento en la siguiente compra por cada usuario que refieran
- T6: Ofrecer 4% de descuento en la siguiente compra por cada usuario que refieran

Te piden ahora medir este experimento (estratificado por `phone_verified`, `population` y `device_value`) en la base `base_evaluacion.Rdata`. Las variables endógenas son `total_purchases_after` que refleja el gasto total post-tratamiento y `transacted` que refleja abrir la app o hacer compras.

Carguemos la base

```
rm(list = ls())

load('Bases input/base_evaluacion.Rdata')

universo_f$treat2<-NULL
```

10 (2ptos). Muestra el estimador ITT para la tasa de transaccionalidad. Recuerda que tu cliente es un grupo empresarial. Por ende, muestra una gráfica donde se aprecie la diferencia entre los grupos de tratamiento y las significancias de manera sencilla. Interpreta tus resultados

- Los descuento variables directos son la forma de incentivo más efectiva.
- En este grupo, incrementar el incentivo 25% (de 20 a 25%) incrementa el impacto mas de 50%. Por ende, se debe explorar el monto al alza un poco más en futuras iteraciones. Los usuarios tienen una elasticidad mayor a 1 a descuentos directos variables.
- Descuentos variables por usuario referido es el segundo grupo de incentivos más efectivo. En este grupo la elasticidad a incrementos de este es menor a 1; por lo que el monto óptimo es menor a 4%.
- Finalmente, los pagos lump sum parecen ser los menos efectivos (si bien igual muestran un lift estadísticamente significativo). En este caso, la elasticidad al monto del cupón muestra una elasticidad menor a 1. Por lo anterior la evidencia arroja que Rappi no debería dar \$200 pesos en lugar de 100.

Para entender de manera integral el sistema de incentivos tenemos que hacer un análisis costo beneficio

```
universo_f<-
  universo_f %>%
  mutate(treat2= case_when(treat ==0 ~ "Control",
                           treat ==1 ~ "Cupón $100",
                           treat ==2 ~ "Cupón $200",
                           treat ==3 ~ "Descuento 20%",
                           treat ==4 ~ "Descuento 25%",
                           treat ==5 ~ "Referidos \n 2%",
                           treat ==6 ~ "Referidos \n 4%"),
         treat2 = factor(treat2, levels = c("Control", "Cupón $100", "Cupón $200",
                                           "Descuento 20%", "Descuento 25%", "Referidos \n 2%", "Referidos \n 4%")))

conteos<-
  universo_f %>%
  group_by(treat2) %>%
  summarise(transacted = mean(transacted),
            total_purchases_after = mean(total_purchases_after))

itt<-impact_eval(data = universo_f,
                 endogenous_vars = c("transacted", "total_purchases_after"),
                 treatment = "treat2", fixed_effect_vars = "strata")

itt$transacted
```

```
# A tibble: 6 x 5
  term                                estimate std.error statistic  p.value
  <chr>                                <dbl>     <dbl>     <dbl>    <dbl>
1 "factor(treat2)Cupón $100"          0.0309    0.0124      2.49 1.27e- 2
2 "factor(treat2)Cupón $200"          0.0444    0.0124      3.58 3.42e- 4
3 "factor(treat2)Descuento 20%"       0.0581    0.0124      4.69 2.72e- 6
4 "factor(treat2)Descuento 25%"       0.0992    0.0124      8.00 1.25e-15
5 "factor(treat2)Referidos \n 2%"     0.0316    0.0124      2.54 1.10e- 2
```



```

6 "factor(treat2)Referidos \n 4%"    0.0705    0.0123    5.73 1.04e- 8
itt<-map(itt, function(x) x %>% mutate(term = str_sub(term, start = 15)))

conteos<-
  left_join(conteos, itt$transacted %>% select(term, p.value), by = c('treat2'='term'))

conteos<-
  conteos %>%
  rename(p.value_transacted = p.value)

conteos<-
  left_join(conteos, itt$total_purchases_after %>% select(term, p.value), by = c('treat2'='term'))

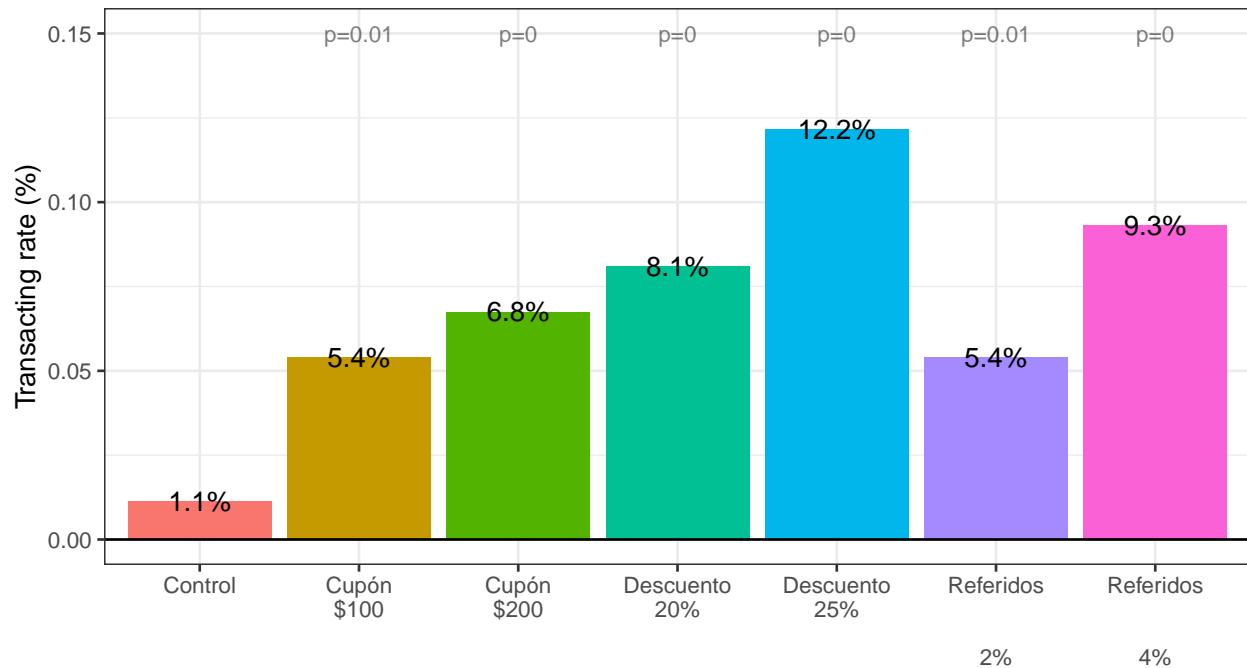
conteos<-
  conteos %>%
  rename(p.value_purchases = p.value)

conteos<-
  conteos %>%
  mutate(p.value_transacted = str_c("p=", round(p.value_transacted, digits = 2)),
         p.value_purchases = str_c("p=", round(p.value_purchases, digits = 2)))

library(scales)
ggplot(conteos, aes(treat2, transacted, label = percent(transacted, accuracy = 0.1)))+
  geom_col(aes(fill = treat2))+theme_bw() + geom_text()+
  labs(y = "Transacting rate (%)", x = " ") +
  scale_x_discrete(labels = function (x) gsub('\\s', '\n', x)) +
  geom_text(aes(treat2, y = 0.15, label = p.value_transacted), color = "black", alpha = 0.5,
            inherit.aes = F, size = 3) + theme(legend.position = "none") +
  geom_hline(yintercept = 0)

```

Warning: Removed 1 rows containing missing values (geom_text).

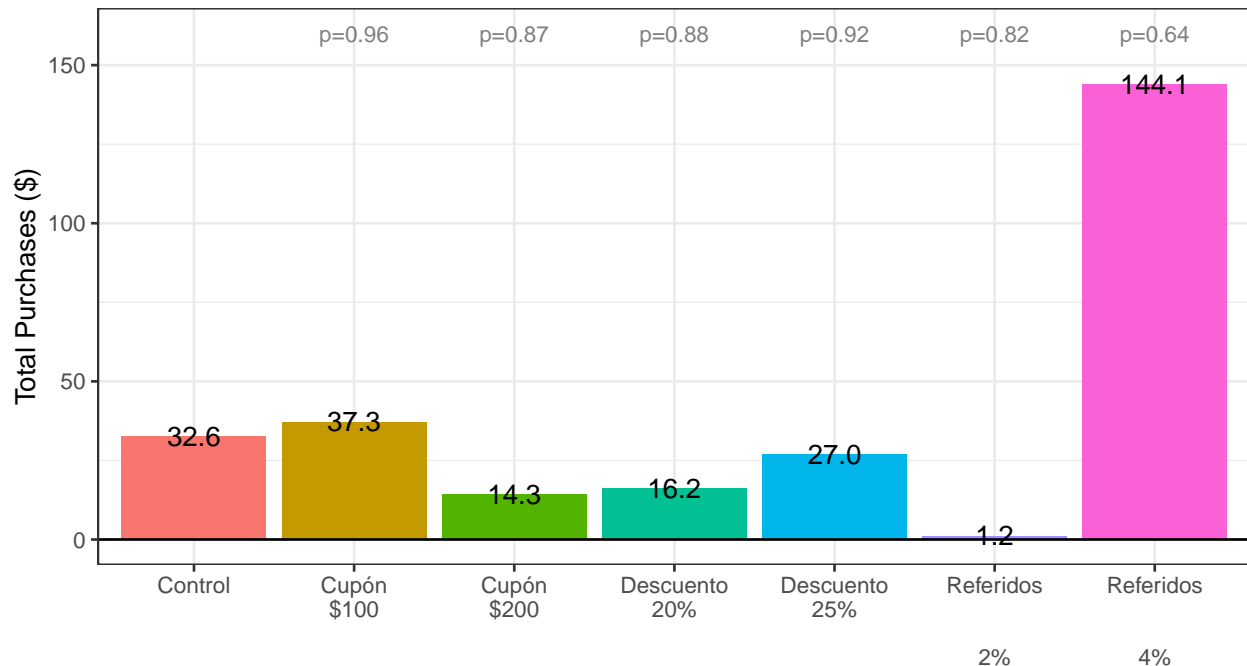


11. Repite el ejercicio sobre compras totales. Que resultados se aprecian? Que indica esto sobre la rentabilidad del sistema de incentivos?

Ninguno de los grupos mostró un lift significativo vs el grupo de control. Esto indica que, en promedio, dar estos incentivos no es efectivo para incrementar las compras en la app de Rappi.

```
ggplot(conteos, aes(treat2, total_purchases_after, label = comma(total_purchases_after)))+
  geom_col(aes(fill = treat2))+theme_bw() + geom_text()+
  labs(y = "Total Purchases ($)", x = " ") +
  scale_x_discrete(labels = function (x) gsub('\\s', '\\n', x)) +
  geom_text(aes(treat2, y = 160, label = p.value_purchases), color = "black", alpha = 0.5,
            inherit.aes = F, size = 3) + theme(legend.position = "none") +
  geom_hline(yintercept = 0)
```

Warning: Removed 1 rows containing missing values (geom_text).



12. Interpreta el impacto del grupo de referidos 4%. Porque el estimador es tan diferente y a la vez es no significativo? Por que esto no paso en la tasa de transaccionalidad?

Esto indica que la varianza del estimador es gigante. Lo anterior se puede leer como una señal ruidosa. Dado que las compras son una variable continua, está muy sujeta a la presencia de outliers. Por ende, los estimadores pueden ser muy ruidosos.

13 (2ptos). Repite la medición en 11 pero ahora con $\log(\text{total_purchases_after}+1)$. Que encuentras ahora? Interpreta las diferencias

Al transformar la variable de compras a logaritmo, encontramos que el descuento de 4% por cada usuario que el tratado refiera incrementa las compras en 15%. Esto muestra que el impacto pudo venir de la tendencia central (i.e. mediana) y que algunos outliers introducían mucho ruido a nuestra medición.

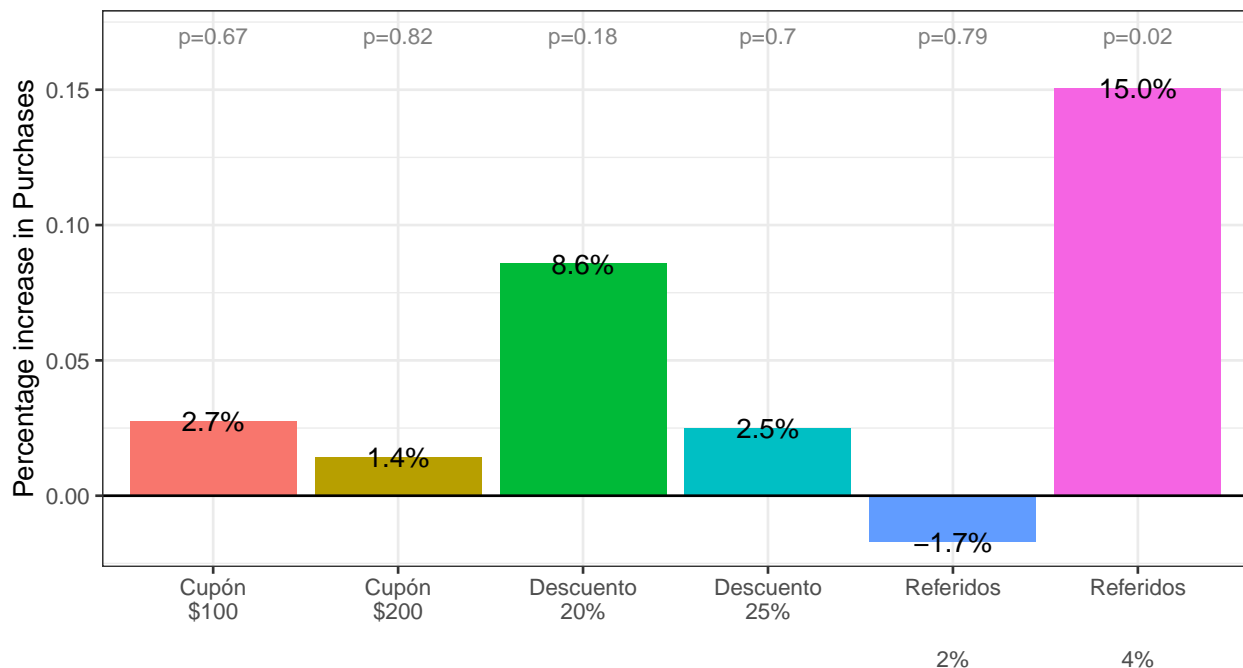
```
universo_f<-
  universo_f %>%
  mutate(log_purchases = log(total_purchases_after +1 ))

itt<-impact_eval(data = universo_f,
                 endogenous_vars = c("log_purchases"),
                 treatment = "treat2", fixed_effect_vars = "strata")

log_purchases<-itt$log_purchases

log_purchases<-
  log_purchases %>%
  mutate(p.value = str_c("p=", round(p.value, digits = 2)),
         term = str_sub(term, start = 15))
```

```
ggplot(log_purchases, aes(term, estimate, label = percent(estimate, accuracy = 0.1)))+
  geom_col(aes(fill = term))+theme_bw() + geom_text()+
  labs(y = "Percentage increase in Purchases", x = " ") +
  scale_x_discrete(labels = function(x) gsub('\\s', '\\n', x)) +
  geom_text(aes(term, y = 0.17, label = p.value), color = "black", alpha = 0.5,
            inherit.aes = F, size = 3) + theme(legend.position = "none") +
  geom_hline(yintercept = 0)
```



14 (4ptos). Describe que variables necesitas para hacer un análisis costo beneficio completo. Les doy algunas: Ticket promedio \$100, Customer Lifetime value: \$1,100. Con esto, que sistema de incentivos recomendarías? Porqué? Muestra el razonamiento detrás de tu recomendación

15 (2ptos). Realiza la estimación de efectos heterogeneos para ambas variables usando population. Que encuentras? existe alguna subpoblación para la que los efectos difieran del promedio? Para cada efecto, muestra gráficas como lo hiciste en los ITTs

Para los usuarios inactivos, los incentivos son más efectivos en general. En particular, dar cupones o descuentos son métodos eficientes. Los descuentos incrementan, en promedio, 15 puntos porcentuales la tasa de transaccionalidad. Por otro lado, para los usuarios nunca activos también se encuentran impactos significativos. No obstante los cambios son mucho más modestos

Por otro lado, se observa que los usuarios nunca activos no muestran ningún impacto en el monto de sus compras al recibir incentivos. El único impacto parece derivarse de los usuarios inactivos que recibieron incentivos de 4% de descuento por usuario referido.

```
itt<-impact_eval(data = universo_f,
  endogenous_vars = c("transacted", "log_purchases"), heterogenous_vars = c("phone_verif",
  treatment = "treat2", fixed_effect_vars = "strata")
```

```

conteos<-
  universo_f %>%
  group_by(treat2, population) %>%
  summarise(transacted = mean(transacted),
            total_purchases_after = mean(total_purchases_after))

itt<-map(itt, function(x) x %>% mutate(term = str_sub(term, start = 15)))
itt<-map(itt, function(x) x %>% mutate(p = str_c('p=',round(p.value, digits = 2))))
itt<-map(itt, function(x) x %>% select(-c(std.error, statistic)))

conteos<-
  left_join(conteos, itt$transacted_population %>% select(term, p, population), by = c('treat2'='term',

conteos<-
  conteos %>%
  rename(p.value_transacted = p)

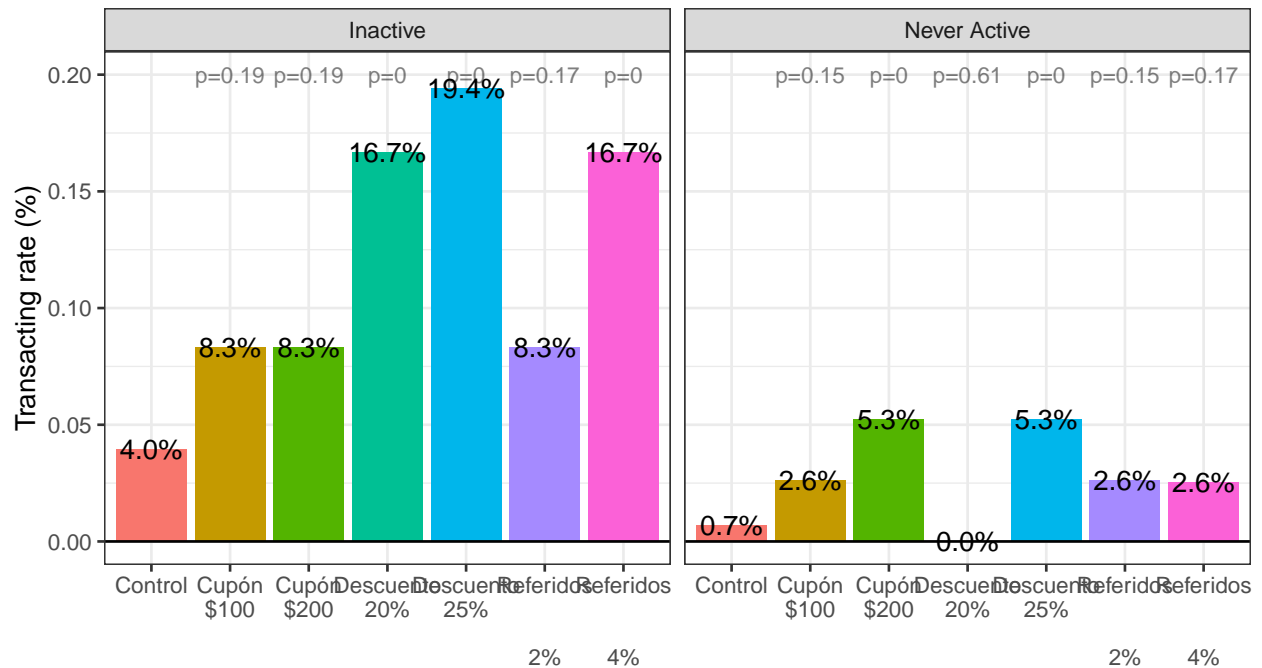
conteos<-
  left_join(conteos, itt$log_purchases_population %>% select(term, p, estimate, population), by = c('tr

conteos<-
  conteos %>%
  rename(p.value_purchases = p)

library(scales)
ggplot(conteos, aes(treat2, transacted, label = percent(transacted, accuracy = 0.1)))+
  geom_col(aes(fill = treat2))+theme_bw() + geom_text()+
  labs(y = "Transacting rate (%)", x = " ") +
  scale_x_discrete(labels = function (x) gsub('\\s', '\\n', x)) +
  geom_text(aes(treat2, y = 0.20, label = p.value_transacted), color = "black", alpha = 0.5,
            inherit.aes = F, size = 3) + theme(legend.position = "none") +
  geom_hline(yintercept = 0)+facet_wrap(~population)

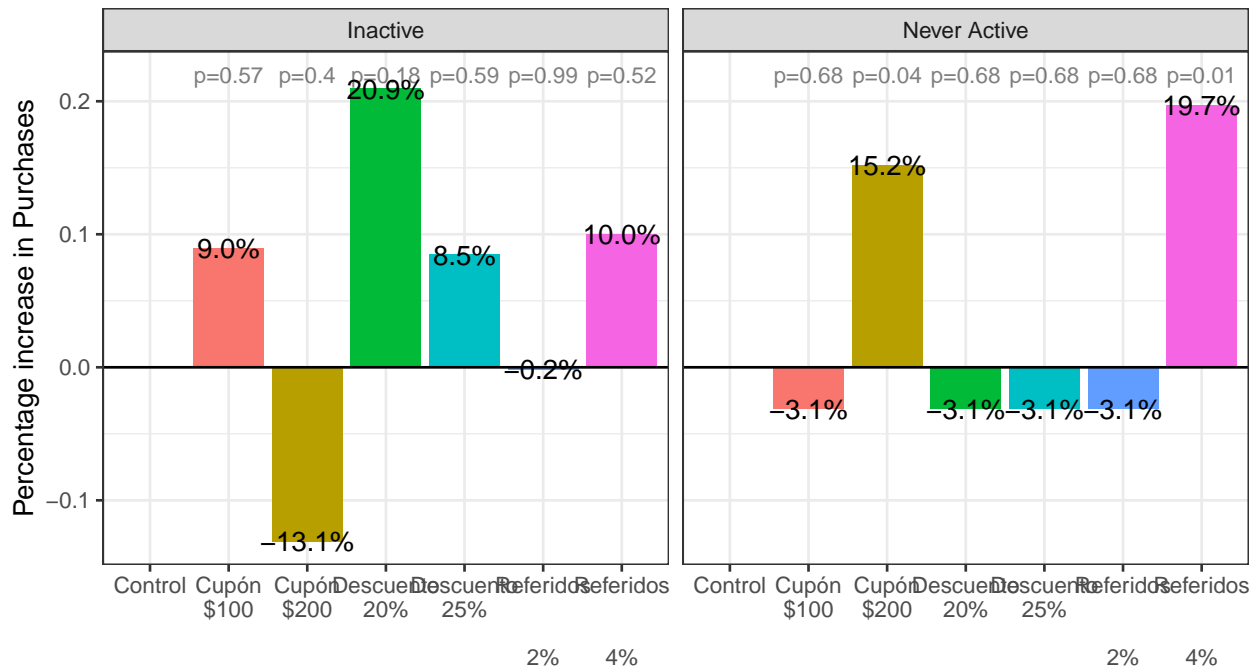
```

Warning: Removed 2 rows containing missing values (geom_text).



```
ggplot(conteos, aes(treat2, estimate, label = percent(estimate, accuracy = 0.1)))+
  geom_col(aes(fill = treat2))+theme_bw() + geom_text()+
  labs(y = "Percentage increase in Purchases", x = " ") +
  scale_x_discrete(labels = function (x) gsub('\s', '\n', x)) +
  geom_text(aes(treat2, y = 0.22, label = p.value_purchases), color = "black", alpha = 0.5,
            inherit.aes = F, size = 3) + theme(legend.position = "none") +
  geom_hline(yintercept = 0)+facet_wrap(~population)
```

Warning: Removed 2 rows containing missing values (position_stack).
 Removed 2 rows containing missing values (geom_text).
 Removed 2 rows containing missing values (geom_text).



16 (2ptos). Repite el ejercicio para 'phone_verified'.

Finalmente, observamos que los usuarios que tenían su teléfono verificado mostraron impactos mucho mayores a los usuarios sin teléfono verificado. Esto se puede deber a que el flujo del usuario tiene más fricción en este caso. Tiene que realizar el flujo de verificación para poder cobrar los incentivos.

```
conteos<-
  universo_f %>%
  group_by(treat2, phone_verified) %>%
  summarise(transacted = mean(transacted),
            total_purchases_after = mean(total_purchases_after))

conteos<-
  left_join(conteos, itt$transacted_phone_verified %>% select(term, p, phone_verified), by = c('treat2', 'phone_verified'))

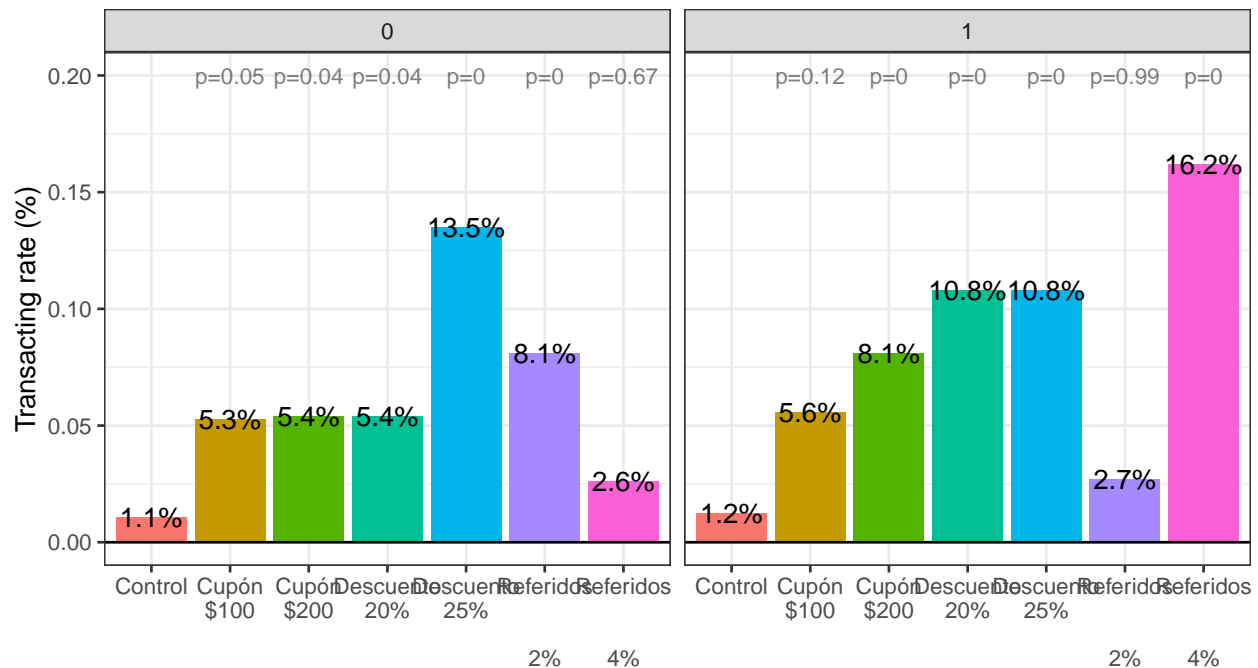
conteos<-
  conteos %>%
  rename(p.value_transacted = p)

conteos<-
  left_join(conteos, itt$log_purchases_phone_verified %>% select(term, p, estimate, phone_verified),
            by = c('treat2'='term', 'phone_verified'))

conteos<-
  conteos %>%
  rename(p.value_purchases = p)
```

```
library(scales)
ggplot(conteos, aes(treat2, transacted, label = percent(transacted, accuracy = 0.1)))+
  geom_col(aes(fill = treat2))+theme_bw() + geom_text()+
  labs(y = "Transacting rate (%)", x = " ") +
  scale_x_discrete(labels = function(x) gsub('\\s', '\n', x)) +
  geom_text(aes(treat2, y = 0.20, label = p.value_transacted), color = "black", alpha = 0.5,
            inherit.aes = F, size = 3) + theme(legend.position = "none") +
  geom_hline(yintercept = 0)+facet_wrap(~phone_verified)
```

Warning: Removed 2 rows containing missing values (geom_text).

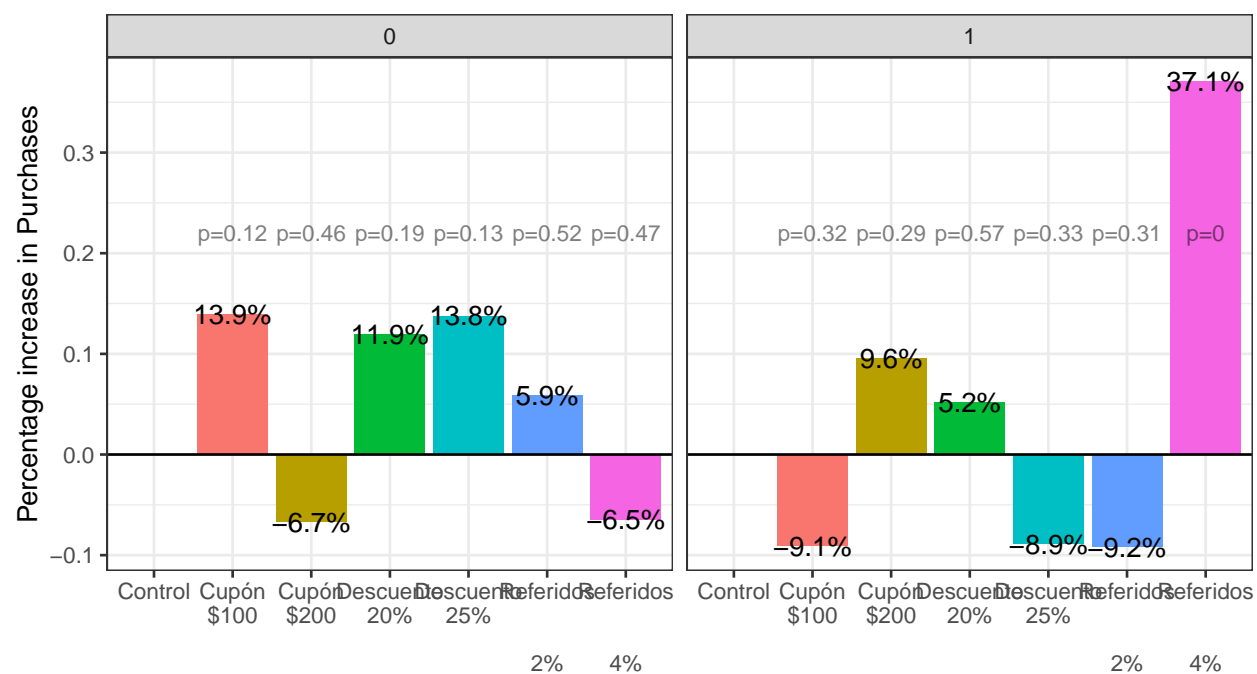


```
ggplot(conteos, aes(treat2, estimate, label = percent(estimate, accuracy = 0.1)))+
  geom_col(aes(fill = treat2))+theme_bw() + geom_text()+
  labs(y = "Percentage increase in Purchases", x = " ") +
  scale_x_discrete(labels = function(x) gsub('\\s', '\n', x)) +
  geom_text(aes(treat2, y = 0.22, label = p.value_purchases), color = "black", alpha = 0.5,
            inherit.aes = F, size = 3) + theme(legend.position = "none") +
  geom_hline(yintercept = 0)+facet_wrap(~phone_verified)
```

Warning: Removed 2 rows containing missing values (position_stack).

Removed 2 rows containing missing values (geom_text).

Removed 2 rows containing missing values (geom_text).



17. Presenta una propuesta de focalización con base en tus resultados generales y heterogeneos.