

Universidad Torcuato Di Tella

DEPARTAMENTO DE ECONOMÍA

FREE VS FEE: UN MODELO DE  
PLATAFORMAS Y DISCRIMINACIÓN DE  
PRECIOS

*Tesis Licenciatura en Economía*

Cufre, Maia - Díaz Geliot, Bautista - Piqueras, Camila - Terfi, Francisco

Tutor: Tappata, Mariano

Agosto 2021

# 1 Introducción

La era de la publicación digital comienza alrededor de 1993 con la consolidación de Internet tal como la conocemos hoy: un mercado masivo, de alcance global, en el cual se ofrece contenido, comunicación y todo tipo de servicios. Los medios de información tradicionales adoptaron, prontamente, estrategias para posicionarse en este nuevo mundo que se asomaba (por ejemplo, NYTimes.com surge en 1996), y ya para 1999 un tercio de los diarios a nivel mundial contaban con ediciones online (Küng, Newman & Picard (2016)). Sin embargo, las versiones digitales eran vistas como una extensión de la versión en papel. Más tarde, con el continuo desarrollo de la tecnología y con el auge de las GAFAN (Google, Apple, Facebook, Amazon y Netflix), las ediciones digitales se colocaron como un elemento clave para los medios periodísticos y no como un anexo al mismo.

A la hora de materializar el valor creado, los diarios digitales comenzaron con una estrategia de financiamiento vía suscripción; sin embargo, a excepción de algunos medios financieros (como *The Wall Street Journal*), esta estrategia fracasó. Frente a esto, comenzaron a explorarse alternativas de financiamiento vía publicidad, dejando el contenido liberado para todo aquel que quiera consumirlo. Sin embargo, para que esta estrategia sea redituable, es necesaria una gran masividad y empresas como Google o Amazon hacen que la competencia por publicidad sea difícil. Es por esta razón que la mayor parte de los medios periodísticos digitales aplican, hoy en día, una estrategia mixta: existe una libre disponibilidad para cierto número de artículos (o notas<sup>1</sup>), mientras que para poder gozar de un consumo sin restricciones es necesario pagar un *fee* en concepto de suscripción al medio. Por tanto, el objetivo de este trabajo consiste en estudiar el *trade-off* que existe entre ingresos por suscripción e ingresos publicitarios, donde el punto clave que los diferencia es la masividad que alcanza el medio (es decir, el número de lecturas).

Para analizar este tópico utilizaremos un modelo teórico con el fin de motivar el comportamiento de los usuarios y la decisión óptima del medio; para luego realizar un estudio empírico sobre datos de un medio digital local y poder contrastar los resultados teóricos.

El modelo teórico comprende el problema de un monopolista (el diario) que debe elegir el precio de suscripción y la cantidad de artículos disponibles sin cargo de forma tal de maximizar sus beneficios. Por el lado de los usuarios, planteamos un continuo de agentes heterogéneos en una única dimensión: la valoración de artículos que lee cada consumidor. El modelo se construye sobre dos ejes claves de la literatura de organización industrial: los mercados multilaterales (o también llamados plataformas) y la discriminación por calidad (o *versioning*).

En Armstrong (2006) y Rochet & Tirole (2003) se define a una plataforma como un intermediario entre dos grupos que interactúan y que generan una externalidad uno sobre el otro. En nuestro caso, los grupos corresponden a los lectores del medio y los publicistas que interactúan a través del diario digital. Ahora bien, un punto que nos distingue de la literatura clásica de mercados multilaterales radica en que nosotros modelamos el mercado publicitario bajo competencia perfecta. Esta simplificación nos permitirá abstraernos del análisis de los publicistas y concentrarnos en el

---

<sup>1</sup>En este trabajo utilizaremos los términos *artículo* o *nota* como sinónimos.

comportamiento de los lectores y la estrategia de la plataforma<sup>2</sup>. Es importante mencionar que la externalidad de los publicistas hacia los lectores está modelada y comprende una externalidad negativa, la exposición a la publicidad genera molestia al lector.

Por otro lado, en nuestro trabajo se encuentran rasgos de *versioning*. En Denecker & McAfee (1996) se argumenta que puede ser óptimo para un monopolista producir un sustituto inferior, de forma tal que los consumidores que valoran el producto superior lo suficiente consuman la versión de baja calidad; los autores sostienen que la forma más barata de hacerlo es *dañando* el producto original. En Varian (1997) se expone que para los denominados *bienes informativos* (léase artículos o notas) una estrategia de pricing basada en el costo no tiene sentido, ya que el costo de producción es muy bajo y existe mucha variabilidad en la valoración por parte de los consumidores; por esta razón, una estrategia como *versioning* cobra mayor relevancia. Lo esencial de esta estrategia radica en que, en un contexto de asimetría informativa, los consumidores se auto-clasifiquen en grupos según su disposición a pagar, de forma tal que la firma puede determinar el precio a partir de una característica endógena (y no observable) como es la elección de calidad.

En nuestro modelo, el monopolista ofrece libre acceso al contenido de la plataforma a cambio de un *fee* (producto de calidad alta); sin embargo, también ofrece una cantidad de artículos fija (máxima) a la cual se puede acceder sin pagar (producto de calidad baja, el bien dañado). La estrategia de liberar cierto número de artículos puede resultar óptima si el precio de la publicidad es suficientemente alto, permitiendo así a la plataforma obtener una mayor masividad y en consecuencia mayores ingresos por publicidad. Con respecto al precio de la suscripción, el monopolista lo determina de forma tal de maximizar sus ganancias, y a diferencia del resultado en McAfee (2007), éste depende, entre otros parámetros, de la distribución de valoraciones por los artículos leídos.

Lo novedoso de este trabajo, además del modelado empleando tanto plataformas multilaterales como discriminación por calidad, viene de la mano del acceso a una base de datos de un medio local que nos permitirá analizar las conclusiones del modelo teórico en un plano empírico. Contamos con datos sobre consumidores suscriptos a la plataforma, lectores que eligen no suscribirse al medio y con datos de un experimento en el cual se liberaron, de forma aleatoria, a una sub-muestra de no suscriptores, garantizándoles libre acceso al contenido. Nuestra hipótesis puede sintetizarse en que los usuarios leen notas porque las mismas les generan utilidad, y por esta razón estarán dispuestos a pagar por una suscripción si su utilidad es lo suficientemente alta. De esta manera, el hecho que paguen no modifica las notas que deciden leer, solo que ahora tienen acceso ilimitado a ese contenido y pueden alcanzar su nivel óptimo (antes leían menos por enfrentar una restricción operativa). Con los datos que contamos, podremos comparar las distribuciones de todos los tipos de lectores y veremos que los resultados del modelo teórico se condicen con la evidencia empírica. Además, intentaremos realizar una primera aproximación a la distribución de valoraciones por artículos (una variable clave en nuestro modelo, pero inobservable en los datos).

Nuestro trabajo se estructura de la siguiente forma:

---

<sup>2</sup>La firma es tomadora de precios en el mercado publicitario, y para cualquier nivel de espacio publicitario que resuelva va a encontrar demandantes de dicho espacio.

- En la sección 2 se desarrolla el modelo teórico previamente mencionado.
- La sección 3 tiene por objeto clarificar y explicar debidamente el funcionamiento institucional y las características particulares del medio local que nos brindó los datos para el trabajo. Por cuestiones de confidencialidad nos reservamos el nombre.
- La sección 4 comprende un exhaustivo análisis de la base de datos (estadísticos descriptivos, distribuciones, entre otros) y un contraste con los resultados teóricos desarrollados en la sección 2.
- En la sección 5 se analiza, en forma particular, el comportamiento de los suscriptores en el plano empírico.
- En la sexta sección analizamos los puntos débiles de nuestro análisis empírico, citando literatura que haya abordado la problemática para tener un proxy de los resultados pendientes.
- Finalmente, en la séptima sección presentamos nuestras conclusiones.
- Los apéndices se encuentran al final del trabajo.

## 2 Modelo Teórico

Suponemos un continuo de consumidores de masa 1, cada uno caracterizado por un parámetro  $\theta$  que captura la valoración por los artículos que lee. Este parámetro es el único grado de heterogeneidad entre los consumidores y permitirá que cada tipo consuma una cantidad distinta de artículos en equilibrio. La utilidad de un consumidor tipo  $\theta$  que lee  $x$  artículos es

$$U(x, s; \theta) = \theta\phi(x) - p \cdot s_\theta - k(x)$$

donde  $\theta\phi(x)$  es la utilidad bruta que recibe de leer  $x$  artículos,  $k(x)$  captura el malestar generado por la exposición a la publicidad,  $p$  es el precio que paga por suscribirse a la plataforma y  $s_\theta$  es una función indicadora tal que

$$s_\theta = \begin{cases} 1 & \text{si el consumidor se suscribe} \\ 0 & \text{caso contrario} \end{cases}$$

Suponemos que  $\phi' > 0$ ,  $\phi'' < 0$ ,  $k' > 0$ ,  $k'' > 0$  y  $\phi(0) = k(0) = 0$ . Además, el parámetro  $\theta$  se distribuye según una función de distribución acumulada  $F(\theta)$  (con densidad  $f(\theta)$ ) y soporte  $[\underline{\theta}, \bar{\theta}]$ . Cada tipo de consumidor enfrenta dos decisiones: si suscribirse o no al medio y la cantidad de artículos que lee (dada su decisión con respecto a la suscripción). Asumimos que ambas decisiones se toman en un contexto estático.

**Proposición 1.** *La decisión socialmente óptima de artículos para un consumidor tipo  $\theta$ ,  $x^*(\theta)$ , es independiente del precio de suscripción  $p$  y es creciente en  $\theta$ .*

Un consumidor tipo  $\theta$  resuelve el siguiente problema

$$\max_x U = \theta\phi(x) - ps - k(x)$$

La condición de primer orden (necesaria y suficiente) viene dada por <sup>3</sup>

$$\theta\phi'(x^*(\theta)) - k'(x^*(\theta)) = 0 \tag{1}$$

De (1) se sigue que

$$\frac{\partial x^*(\theta)}{\partial \theta} = -\frac{\phi'(x^*(\theta))}{\theta\phi''(x^*(\theta)) - k''(x^*(\theta))} > 0$$

Asumimos que existe un tipo de consumidor (que denotaremos como  $\theta_1 \in [\underline{\theta}, \bar{\theta})$ ) que está indiferente entre leer y no leer; es decir,

$$\theta_1\phi(x^*(\theta_1)) - k(x^*(\theta_1)) = 0 \tag{2}$$

Notar que  $\theta_1 = \theta_1(k(\cdot))$  y la masa de no lectores es  $F(\theta_1)$ . La decisión de los no lectores es trivial e independiente de la política de la plataforma: eligen no suscribirse ( $s_\theta = 0$ ) y no leen artículos ( $x(\theta) = 0$ ).

---

<sup>3</sup>Además, sabemos que el valor que resuelve (1) implica un máximo para el problema del agente dado que la función objetivo es cóncava.

En lo que refiere al comportamiento de la plataforma, suponemos un monopolista que opera con costos marginales constantes y normalizados a cero y costos fijos nulos. Sus ingresos provienen, potencialmente, de dos fuentes distintas:

1. **Ingresos por suscripción.** Cada suscriptor paga un precio  $p$ , independiente de la cantidad de notas que lee, para tener libre acceso a los artículos del medio.
2. **Ingresos por publicidad.** La firma recibe un ingreso  $q$  por nota leída en concepto de publicidad. Asumimos que el mercado de publicidad es competitivo, por lo que la plataforma es tomadora de precios ( $q$  exógeno).

La plataforma definirá su política (o estrategia) óptima de forma tal de maximizar sus beneficios. Esto implica que la plataforma cobrará un precio  $p$  por suscripción y determinará un cantidad de notas  $\bar{x}$  disponibles para leer sin necesidad de estar suscripto.

## 2.1 Escenario First Best: Discriminación perfecta

A modo de *benchmark* suponemos un escenario con información perfecta en el cual la firma puede cobrar precios diferenciados y ofrecer una cantidad diferenciada de artículos para cada consumidor.

**Definición 1.** *Un equilibrio en un escenario de First Best comprende un vector  $(s_\theta, x(\theta))$ <sup>4</sup> para todo  $\theta \in [\underline{\theta}, \bar{\theta}]$  y una política de la plataforma  $(p(\theta), \bar{x}(\theta))$  tales que*

1. *Cada tipo de consumidor maximiza su utilidad, dada la política de la plataforma.*
2. *La plataforma elige su política de forma tal de maximizar sus beneficios.*

La plataforma ofrecerá a cada tipo de consumidor una cantidad de artículos sin cargo igual a  $\bar{x}(\theta) = 0$  y cobrará a cada uno un precio  $p^{FB}(\theta)$  tal que para todo  $\theta \in [\theta_1, \bar{\theta}]$

$$U(x^*(\theta), 1; \theta) = \theta\phi(x^*(\theta)) - k(x^*(\theta)) - p(\theta) = 0$$

Por lo que

$$p(\theta)^{FB} = \theta\phi(x^*(\theta)) - k(x^*(\theta)) \tag{3}$$

Tenemos un equilibrio en el cual todos los lectores con  $\theta \in [\theta_1, \bar{\theta}]$  se suscriben a la plataforma y cada uno lee su cantidad de artículos óptima. La plataforma extrae todo el excedente.

Los beneficios de la plataforma serán<sup>5</sup>:

$$\Pi^{FB} = \int_{\theta_1}^{\bar{\theta}} [\theta\phi(x^*(\theta)) - k(x^*(\theta)) + qx^*(\theta)] dF(\theta) \tag{4}$$

Es importante resaltar que este escenario es eficiente, en el sentido de que se maximiza el bienestar agregado.

---

<sup>4</sup>Este vector corresponde a las decisiones que toma cada tipo de consumidor: suscribirse o a la plataforma y número de artículos que lee.

<sup>5</sup>El primer sumando son los ingresos por suscripción y el segundo por publicidad.

## 2.2 Escenario Second Best: Asimetría informativa

Consideremos ahora el caso en el cual el tipo de cada individuo es información privada, por lo que la plataforma no puede condicionar la cantidad de artículos ni el precio de suscripción a esta característica.

**Definición 2.** *Un equilibrio en un escenario de Second Best con asimetría informativa comprende un vector  $(s_\theta, x(\theta))$  para todo  $\theta \in [\underline{\theta}, \bar{\theta}]$  y una política de la plataforma  $(p, \bar{x})$  tales que:*

1. *Cada tipo de consumidor maximiza su utilidad, dada la política de la plataforma.*
2. *La plataforma elige su política de forma tal de maximizar sus beneficios.*

En este marco, existen tres casos que resultan interesantes para nuestro análisis.

### Caso 1: Escenario sin suscripción y sin restricciones de lectura

Éste escenario comprende un caso en el cual la plataforma busca maximizar los ingresos por publicidad. Para que esto ocurra, la misma no debe cobrar por suscripción y debe haber libre disponibilidad de artículos para todos los consumidores. Luego, la política de la plataforma en equilibrio es  $p_1^{SB} = 0$  y/o  $\bar{x}_1^{SB} \geq x^*(\bar{\theta})$ .

Dada esta política, la decisión de suscripción a la plataforma por parte de los consumidores no es relevante puesto que no existe costo para suscribirse. Con respecto a la decisión de lectura, cada consumidor lee la cantidad de artículos socialmente óptima,  $x^*(\theta)$ , ya que la restricción de lectura no es operativa.

En este caso, por cada tipo de consumidor la plataforma recibe un ingreso proveniente únicamente de la publicidad a la que se expone cada lector, es decir  $\pi_1^{SB}(\theta) = qx^*(\theta)$ . Luego, los beneficios que obtiene la plataforma si decide seguir esta estrategia vienen dados por

$$\Pi_1^{SB} = \int_{\theta_1}^{\bar{\theta}} qx^*(\theta)dF(\theta) \quad (5)$$

Bajo esta política el excedente del consumidor es

$$EC_1^{SB} = \int_{\theta_1}^{\bar{\theta}} [\theta\phi(x^*(\theta)) - k(x^*(\theta))]dF(\theta) \quad (6)$$

Es importante resaltar que en este caso se maximiza el excedente agregado, ya que

$$\Pi^{FB} = \Pi_1^{SB} + EC_1^{SB}$$

es decir, la asimetría informativa no tiene costo en términos de eficiencia.

## Caso 2: Lectura solo con suscripción

Consideremos ahora otro caso extremo en el cual la plataforma solo permite leer a quiénes decidan suscribirse, pagando un precio  $p_2^{SB}$ . Todo consumidor que elige suscribirse tiene libre acceso a los artículos que publica la plataforma, y por lo tanto cada suscriptor leerá el número de notas que maximiza su utilidad.

Dada esta política, existe un consumidor tipo  $\hat{\theta} \in [\underline{\theta}, \bar{\theta}]$  tal que está indiferente entre suscribirse y no hacerlo

$$\hat{\theta}\phi(x^*(\hat{\theta})) - p_2^{SB} - k(x^*(\hat{\theta})) = 0 \quad (7)$$

Luego, para todo  $\theta \in [\underline{\theta}, \hat{\theta})$  la decisión es no suscribirse ( $s_\theta = 0$ ) y por consiguiente no leer notas ( $x(\theta) = 0$ ); mientras que para todo  $\theta \in [\hat{\theta}, \bar{\theta}]$  la decisión es suscribirse ( $s_\theta = 1$ ) y por lo tanto leer la cantidad óptima de artículos ( $x(\theta) = x^*(\theta)$ ).

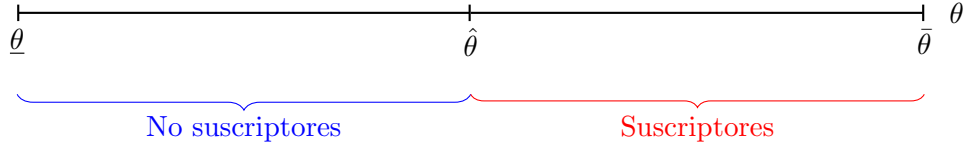


Figure 1: Tipología de consumidores en caso 2

La masa de suscriptores es  $S_2^{SB} = 1 - F(\hat{\theta})$  y depende de la política de la plataforma, ya que  $\hat{\theta}$  depende del precio de suscripción. En este contexto imponemos que  $\bar{x} = 0$  (para poder leer hay que suscribirse) y la plataforma elige el precio  $p$  tal que maximiza sus beneficios

$$\max_p \Pi = \int_{\hat{\theta}}^{\bar{\theta}} [p + qx^*(\theta)] dF(\theta) \quad (8)$$

La condición de primer orden del problema (necesaria y suficiente) es

$$\begin{aligned} \frac{\partial \Pi}{\partial p} &= \frac{\partial F(\hat{\theta})}{\partial \hat{\theta}} \frac{\partial \hat{\theta}}{\partial p} - \frac{\partial \hat{\theta}}{\partial p} [p + qx^*(\hat{\theta})] \frac{\partial F(\hat{\theta})}{\partial \hat{\theta}} + \int_{\hat{\theta}}^{\bar{\theta}} dF(\theta) = 0 \\ &= \frac{\partial \hat{\theta}}{\partial p} \frac{\partial F(\hat{\theta})}{\partial \hat{\theta}} [q(\bar{x} - x^*(\hat{\theta})) - p] + (1 - F(\hat{\theta})) = 0 \end{aligned} \quad (9)$$

donde la ecuación (9) caracteriza la elección de  $p_2^{SB}$ .

Luego, un suscriptor tipo  $\theta$  brinda un ingreso  $\pi_2^{SB}(\theta) = p_2^{SB} + qx^*(\theta)$  al medio; por lo que los beneficios de la plataforma en este caso vienen dados por

$$\Pi_2^{SB} = \int_{\hat{\theta}}^{\bar{\theta}} [p_2^{SB} + qx^*(\theta)] dF(\theta) \quad (10)$$

y el excedente del consumidor es

$$EC_2^{SB} = \int_{\hat{\theta}}^{\bar{\theta}} [\theta\phi(x^*(\theta)) - k(x^*(\theta)) - p_2^{SB}] dF(\theta) \quad (11)$$



Notemos que  $\hat{\theta} > \theta_1$ , ya que

$$\begin{aligned}\theta_1\phi(x^*(\theta_1)) - k(x^*(\theta_1)) &= 0 \\ \hat{\theta}\phi(x^*(\hat{\theta})) - k(x^*(\hat{\theta})) - p_2^{SB} &= 0\end{aligned}$$

Por lo que, dado  $p_2^{SB} > 0$

$$\begin{aligned}\theta_1\phi(x^*(\hat{\theta})) - k(x^*(\hat{\theta})) &< 0 < \hat{\theta}\phi(x^*(\hat{\theta})) - k(x^*(\hat{\theta})) \\ \theta_1 &< \hat{\theta}\end{aligned}$$

Al respecto, es importante resaltar los siguientes puntos:

- Si un escenario como el del caso 2 consistuye un equilibrio, éste sería ineficiente. Los tipos tales que  $\theta \in [\theta_1, \hat{\theta})$  tienen una valoración positiva por los artículos (que se producen a costo cero) y no leen en equilibrio. Esta ineficiencia tiene dos aristas: (1) la **pérdida de bienestar** por parte de estos consumidores que no demandan artículos y; (2) la **pérdida de ingresos publicitarios** de la plataforma. *Ergo*, podemos computar la pérdida de eficiencia de la siguiente forma

$$DWL_2^{SB} = \underbrace{\int_{\theta_1}^{\hat{\theta}} [\theta\phi(x^*(\theta)) - k(x^*(\theta))]dF(\theta)}_{\text{Pérdida de bienestar de lectores}} + \underbrace{\int_{\theta_1}^{\hat{\theta}} qx^*(\theta)dF(\theta)}_{\text{Pérdida de ingresos publicitarios}} \quad (12)$$

- Un escenario como el previamente descrito puede no constituir un equilibrio si  $\bar{x} = 0$  no es óptimo para la plataforma. Es decir, puede ser óptimo para la plataforma elegir un  $\bar{x} > 0$  de forma tal que los consumidores que eligen no suscribirse tengan acceso a notas sin cargo y así generar mayores ingresos publicitarios. Esta situación es la que se analiza en el caso 3 a continuación.

### Caso 3: Escenario con suscripción y con restricción de lectura

Por último, consideremos el caso en el cual la plataforma cobra un precio  $p > 0$  y provee una cantidad de artículos  $\bar{x} \in (0, x^*(\bar{\theta}))$  sin cargo (solución interior). Dada esta política, es importante caracterizar los siguientes *consumidores críticos*:

1. Tipo que le es óptimo leer la cantidad de notas restringida: Existe  $\theta_2$  tal que

$$\theta_2\phi(x_2^*(\theta_2)) - k(x^*(\theta_2)) = \theta_2\phi(\bar{x}) - k(\bar{x}) \quad (13)$$

Notar que  $\theta_2 = \theta_2(k(\cdot), \bar{x})^6$  y la masa de lectores no suscriptos para los cuales la restricción de lectura no es operativa es  $F(\theta_2) - F(\theta_1)$ . De ahora en más llamaremos **esporádico** a un lector tipo  $\theta$  con  $\theta \in [\theta_1, \theta_2]$ . Los lectores esporádicos leen la cantidad de notas óptimas  $x^*(\theta)$ . Recordemos que los consumidores con  $\theta \in [\underline{\theta}, \theta_1)$  son **no lectores**.

---

<sup>6</sup>Es importante tener en cuenta que para el consumidor tipo  $\theta_2$  es óptimo  $x^*(\theta_2) = \bar{x}$  (la función de utilidad del agente es inyectiva).

2. Tipo indiferente entre suscribirse y no hacerlo: Existe  $\theta_3$  tal que

$$\theta_3\phi(x^*(\theta_3)) - p - k(x^*(\theta_3)) = \theta_3\phi(\bar{x}) - k(\bar{x}) \quad (14)$$

Notar que  $\theta_3 = \theta_3(k(\cdot), \bar{x}, p)$  y la masa de **suscriptores** es  $1 - F(\theta_3)$ . Los suscriptores son lectores que leen su cantidad óptima  $x^*(\theta)$ , con  $\theta \in (\theta_3, \bar{\theta}]$ .

Existen lectores de tipo  $\theta \in (\theta_2, \theta_3]$  que no se suscriben, pero para los cuales la restricción de lectura es operativa, de modo que leen  $\bar{x}$  artículos; los denominaremos **potenciales suscriptores**<sup>7</sup>.

El siguiente diagrama resume las tipologías de consumidores en este caso:

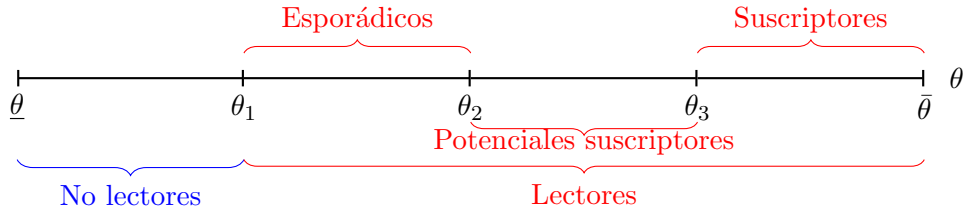


Figure 2: Tipología de consumidores en caso 3

**Proposición 2.** *La masa de suscriptores,  $S_3^{SB}$ , disminuye con  $\bar{x}$  y disminuye con  $p$ .*<sup>8</sup>

De la ecuación (14) se sigue que

$$\frac{\partial \theta_3}{\partial \bar{x}} = \frac{\theta_3 \phi'(\bar{x}) - k'(\bar{x})}{\phi(x^*(\theta_3)) - \phi(\bar{x})} > 0$$

$$\frac{\partial \theta_3}{\partial p} = \frac{1}{\phi(x^*(\theta_3)) - \phi(\bar{x})} > 0$$

Dado que la masa de suscriptores a la plataforma es  $S_3^{SB} = 1 - F(\theta_3)$ , un aumento en la cantidad de artículos disponibles sin cargo ( $\bar{x}$ ) induce un aumento de  $\theta_3$ , disminuyendo la cantidad de suscriptores. Un aumento en  $\bar{x}$  reduce los incentivos a incurrir en el costo de suscripción. A su vez, un aumento en el costo de suscripción reduce, naturalmente, la masa de suscriptores.

Por otro lado, la plataforma determina su política comercial óptima  $(p_3^{SB}, \bar{x}_3^{SB})$  de forma tal de maximizar sus beneficios; por lo que resuelve el siguiente problema

$$\max_{\{p, \bar{x}\}} \Pi = \underbrace{\int_{\theta_1}^{\theta_2} qx^*(\theta) dF(\theta)}_{\text{lectores esporádicos}} + \underbrace{\int_{\theta_2}^{\theta_3} q\bar{x} dF(\theta)}_{\text{suscriptores potenciales}} + \underbrace{\int_{\theta_3}^{\bar{\theta}} (p + qx^*(\theta)) dF(\theta)}_{\text{suscriptores}} \quad (15)$$

<sup>7</sup>Este grupo de agentes será la fuente de ineficiencia del equilibrio en el caso 3.

<sup>8</sup>Para la demostración de esta proposición ver Apéndice 1.

Recordando de qué depende cada  $\theta$  crítico:

$$\max_{\{p, \bar{x}\}} \Pi = \int_{\theta_1}^{\theta_2(\bar{x})} qx^*(\theta) dF(\theta) + q\bar{x} [F(\theta_3(\bar{x}, p)) - F(\theta_2(\bar{x}))] + \int_{\theta_3(\bar{x}, p)}^{\bar{\theta}} (p + qx^*(\theta)) dF(\theta)$$

Las condiciones de primer orden del problema (necesarias y suficientes) vienen dadas por:

$$\begin{aligned} \frac{\partial \Pi}{\partial p} &= \frac{\partial F(\theta_3)}{\partial \theta_3} \frac{\partial \theta_3}{\partial p} - \frac{\partial \theta_3}{\partial p} [p + qx^*(\theta_3)] \frac{\partial F(\theta_3)}{\partial \theta_3} + \int_{\theta_3}^{\bar{\theta}} dF(\theta) = 0 \\ &= \frac{\partial \theta_3}{\partial p} \frac{\partial F(\theta_3)}{\partial \theta_3} [q(\bar{x} - x^*(\theta_3)) - p] + (1 - F(\theta_3)) = 0 \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{\partial \Pi}{\partial \bar{x}} &= \frac{\partial \theta_2}{\partial \bar{x}} qx^*(\theta_2) \frac{\partial F(\theta_2)}{\partial \theta_2} + \left[ \frac{\partial F(\theta_3)}{\partial \theta_3} \frac{\partial \theta_3}{\partial \bar{x}} - \frac{\partial F(\theta_2)}{\partial \theta_2} \frac{\partial \theta_2}{\partial \bar{x}} \right] q\bar{x} + (F(\theta_3) - F(\theta_2))q - \frac{\partial \theta_3}{\partial \bar{x}} (p + qx^*(\theta_3)) \frac{F(\theta_3)}{\partial \theta_3} = 0 \\ &= \frac{\partial \theta_2}{\partial \bar{x}} q(x^*(\theta_2) - \bar{x}) \frac{\partial F(\theta_2)}{\partial \theta_2} - \frac{\partial \theta_3}{\partial \bar{x}} [q(x^*(\theta_3) - \bar{x}) + p] \frac{\partial F(\theta_3)}{\partial \theta_3} = 0 \end{aligned} \quad (17)$$

Las ecuaciones (16) y (17) determinan la elección óptima de artículos disponibles sin cargo y el precio de suscripción en este caso:  $p_3^{SB}$  y  $\bar{x}_3^{SB}$ .<sup>9</sup>

En el siguiente gráfico se muestra la relación entre la cantidad de notas leídas y el parámetro  $\theta$  en un escenario con suscripción y con restricción de lectura (en rojo); y la función  $x^*(\theta)$  que muestra la cantidad de notas óptimas para cada tipo de consumidor (en azul).<sup>10</sup>

<sup>9</sup>En el caso 3 asumimos solución interior al problema del monopolista, es decir  $p > 0$  y  $\bar{x} \in (0, x^*(\bar{\theta}))$ . Las posibles soluciones de esquina están contempladas en los casos 1 y 2, analizados previamente.

<sup>10</sup>Notar que la curva  $x^*(\theta)$  es creciente por la Proposición 1 y es cóncava porque  $\phi(\cdot)$  es cóncava y  $k(\cdot)$  es convexa. El gráfico se realizó para una función estrictamente cóncava, pero es posible obtener, con los supuestos del modelo, una función  $x^*(\theta)$  lineal.

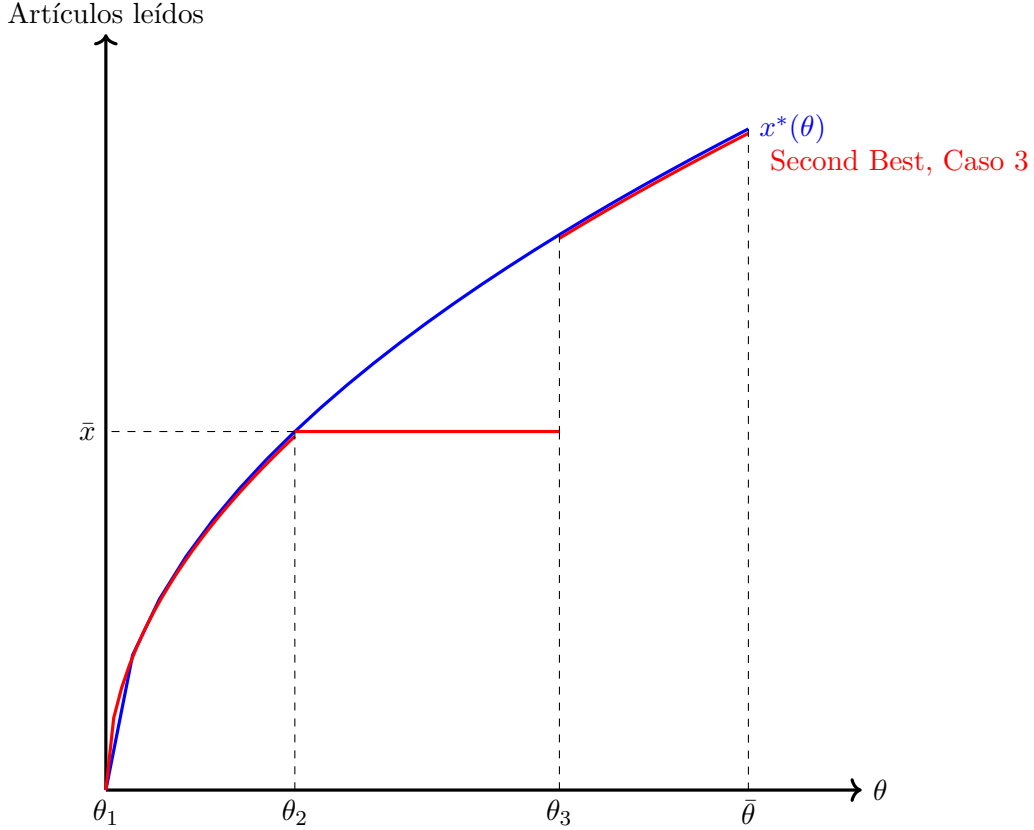


Figure 3: Artículos leídos en función de  $\theta$  en Second Best (Caso 3) y función  $x^*(\theta)$

Notemos que ambas curvas coinciden para los lectores esporádicos,  $\theta \in [\theta_1, \theta_2]$ , y para los lectores que están suscriptos a la plataforma,  $\theta \in (\theta_3, \bar{\theta}]$ . Esto ocurre porque en el caso 3 de un escenario de Second Best ambas tipologías de consumidores leen la cantidad de artículos socialmente óptima, ya sea porque la restricción de lectura no es operativa (lectores esporádicos) o porque incurren en el costo de suscripción y tienen libre acceso a los artículos. Por otro lado, para los potenciales suscriptores,  $\theta \in (\theta_2, \theta_3]$ , la restricción de lectura es operativa por lo que esta tipología de agentes consumen una cantidad ineficientemente baja de artículos cuando existe esta asimetría informativa y la plataforma emplea una política comercial de estas características.

Con respecto a los beneficios de la plataforma, un consumidor tipo  $\theta$  reporta el siguiente beneficio para el monopolista

$$\pi(\theta) = \begin{cases} qx^*(\theta) & \text{si } \theta \in [\theta_1, \theta_2) \\ q\bar{x}_3^{SB} & \text{si } \theta \in [\theta_2, \theta_3) \\ p_3^{SB} + qx^*(\theta) & \text{si } \theta \in [\theta_3, \bar{\theta}] \end{cases}$$

Luego, si la plataforma adopta esta estrategia sus beneficios serán

$$\Pi_3^{SB} = \int_{\theta_1}^{\theta_2} qx^*(\theta)dF(\theta) + \int_{\theta_2}^{\theta_3} q\bar{x}_3^{SB}dF(\theta) + \int_{\theta_3}^{\bar{\theta}} [p_3^{SB} + qx^*(\theta)]dF(\theta) \quad (18)$$

Además, el excedente del consumidor viene dado por

$$EC_3^{SB} = \int_{\theta_1}^{\theta_2} [\theta \phi(x^*(\theta)) - k(x^*(\theta))] dF(\theta) + \int_{\theta_2}^{\theta_3} [\theta \phi(\bar{x}_3^{SB}) - k(\bar{x})] dF(\theta) + \int_{\theta_3}^{\bar{\theta}} [\theta \phi(x^*(\theta)) - p_3^{SB} - k(x^*(\theta))] dF(\theta) \quad (19)$$

Dado que la restricción de lectura es operativa para algunos consumidores, el equilibrio en este escenario será ineficiente. Al igual que ocurría en un equilibrio solo leía el suscriptor, la ineficiencia en este caso tiene dos aristas: (1) por un lado, la **pérdida de bienestar** de estos lectores por leer menos notas y por otro lado; (2) la **pérdida de ingresos publicitarios** para la plataforma debido a que estos lectores leen menos artículos. Luego, podemos computar la pérdida de eficiencia en este caso como

$$DWL_3^{SB} = \underbrace{\int_{\theta_2}^{\theta_3} [\theta \cdot (\phi(x^*(\theta)) - \phi(\bar{x})) - (k(x^*(\theta)) - k(\bar{x}))] dF(\theta)}_{\text{Pérdida de bienestar de lectores}} + \underbrace{\int_{\theta_2}^{\theta_3} [x^*(\theta) - \bar{x}] \cdot q dF(\theta)}_{\text{Pérdida de ingresos publicitarios}} \quad (20)$$

El siguiente cuadro resume los principales resultados de los tres casos previamente expuestos

Caso	Política comercial	Lectores	Suscriptores	Ingresos	¿Eficiente?
1	$p_1^{SB} = 0 \cap \bar{x}_1^{SB} \geq x^*(\bar{\theta})$	$[\theta_1, \bar{\theta}]$	-	Solo publicidad	Sí
2	$p_2^{SB} > 0 \cap \bar{x}_2^{SB} = 0$	$[\hat{\theta}, \bar{\theta}]$	$[\hat{\theta}, \bar{\theta}]$	Publicidad y suscripción	No
3	$p_3^{SB} > 0 \cap \bar{x}_3^{SB} \in (0, x^*(\bar{\theta}))$	$[\theta_1, \bar{\theta}]$	$[\theta_3, \bar{\theta}]$	Publicidad y suscripción	No

La política comercial que adopte la plataforma determinará el tipo de equilibrio en un contexto de Second Best con información asimétrica. Naturalmente, la plataforma adoptará la estrategia que brinde mayores beneficios y esto depende directamente de la distribución de tipos  $F(\theta)$  y el valor del parámetro  $q$ . Por ejemplo, si el valor de  $q$  es *muy alto*, la plataforma tenderá a maximizar las lecturas para así obtener mayores ingresos por sector publicitario resignando ingresos por suscripción.

Por otro lado, si la función de densidad  $f(\theta)$  es asimétrica, y en particular acumula una mayor masa para *valores altos* de  $\theta$  (cerca de  $\bar{\theta}$ ), entonces la plataforma tiene incentivos a aumentar el precio de suscripción, resignar lectores y por lo tanto disminuir los ingresos publicitarios. Queda en evidencia el *trade-off* que enfrenta la firma: masividad de audiencia para explotar los ingresos publicitarios versus restringir el número de lecturas y lograr mayores ingresos vía suscripción al medio.

La comparación entre los casos 1 y 3 resulta principalmente atractiva, ya que contamos con datos del diario digital que analizaremos en la sección empírica, que nos permitirán realizar cálculos aproximados sobre los beneficios y evaluar ambas estrategias. Consideremos el siguiente gráfico a modo ilustrativo en el cual se grafica los beneficios por tipo de consumidor de los casos 1 y 3,<sup>11</sup> es decir

$$\pi_1^{SB}(\theta) = qx^*(\theta) \text{ con } \theta \in [\theta_1, \bar{\theta}]$$

<sup>11</sup>A modo de simplificación asumimos  $x^*(\theta)$  lineal, sin embargo si se tratara de una función estrictamente cóncava el análisis sería análogo.

$$\pi_3^{SB}(\theta) = \begin{cases} qx^*(\theta) & \text{si } \theta \in [\theta_1, \theta_2) \\ q\bar{x}_3^{SB} & \text{si } \theta \in [\theta_2, \theta_3) \\ p_3^{SB} + qx^*(\theta) & \text{si } \theta \in [\theta_3, \bar{\theta}] \end{cases}$$

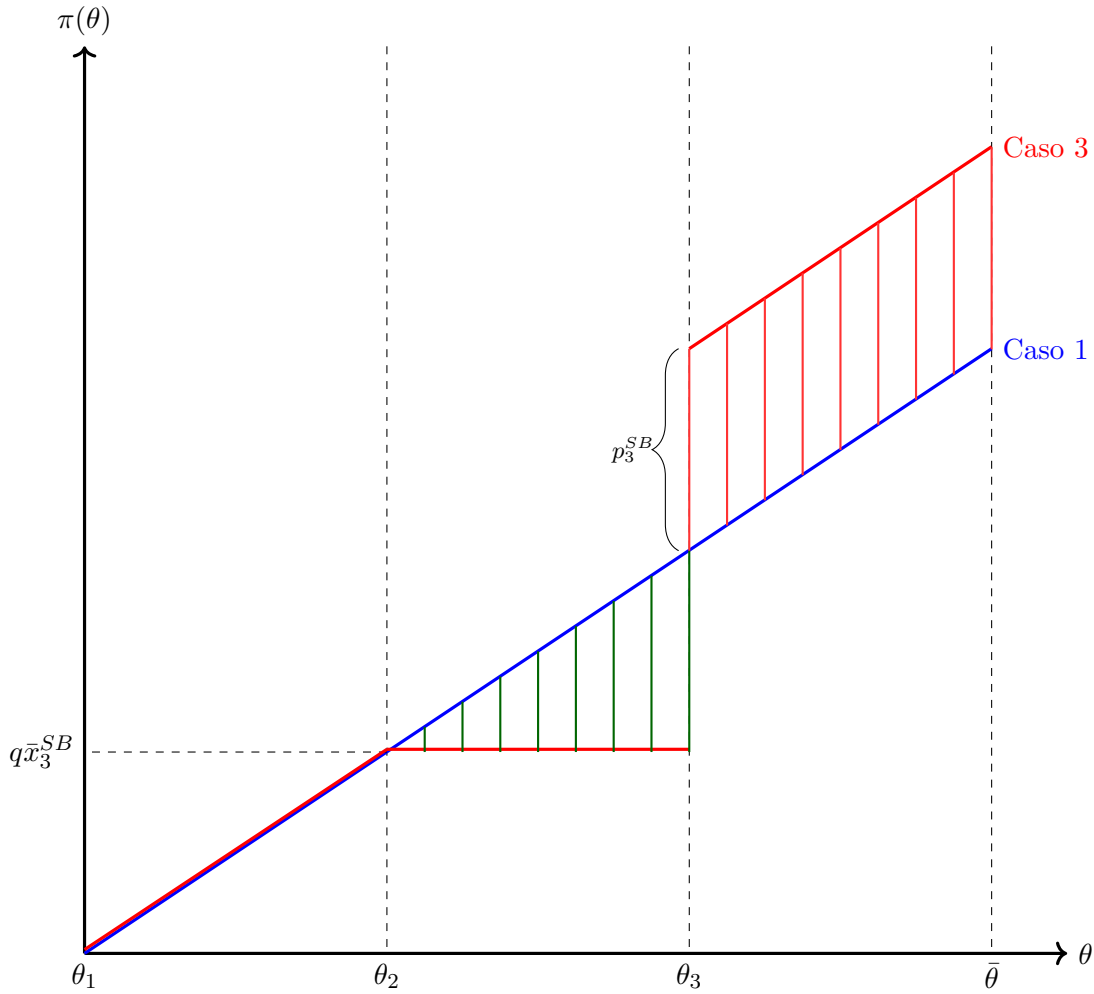


Figure 4: Trade - off suscripción vs publicidad, casos 1 y 3

La plataforma comparará las áreas del triángulo verde y el paralelogramo rosa (ponderando como indique la distribución  $F(\theta)$ , respectivamente). El triángulo comprende los ingresos por publicidad adicionales que se obtienen por liberar las notas en comparación con restringir el número de artículos que se pueden leer sin cargo; mientras que el paralelogramo se condice con los ingresos que provienen de cobrar una tarifa positiva y así obtener libre acceso al material que publica la plataforma.

### 3 La plataforma

Para el análisis empírico del trabajo, contamos con una base de datos de uno de los principales y más tradicionales medios de comunicación de la región. Esta plataforma ha ido modificando su forma de operar en el mercado de la mano de la revolución digital con el fin de adaptarse de la mejor manera posible a las nuevas dinámicas del mercado. A modo de motivación, vamos a realizar una descripción de su modelo de negocios para entender el funcionar de la firma.

En primer lugar, cabe aclarar que estamos trabajando con un medio de comunicación donde su rubro más fuerte son las noticias, tanto en formato de diario de papel como su versión digital. Esta plataforma enfrenta un trade off muy claro entre venta de contenido (suscripción) y venta por espacio publicitario (por nota leída).

Con el mundo digital, las grandes firmas globales migraron a sitios gratuitos que monetizaban sus productos a través de la publicidad. Entre los casos más emblemáticos y exitosos podríamos nombrar a Google, Facebook, Amazon, y para un caso nacional a Mercado Libre (MELI). Sería una percepción muy *naïve* creer que estos ejemplos son análogos a los de nuestra firma de interés, ya que existe una diferencia fundamental: la masividad de los productos que ofrecen, por un lado, pero principalmente que son los propios usuarios los generadores de contenido. Este último punto es primordial, ya que los lectores solo estarán dispuestos a pagar en la medida que se les ofrezcan noticias de calidad (necesitan una diferenciación vertical relativo a las noticias que son de libre acceso en internet).

Llegamos así a nuestro primer punto importante: la firma como entidad (dada su cultura institucional) no está dispuesta a bajar la calidad de las noticias como mecanismo hacia la masividad alcanzando una mayor parte del mercado (bajar el costo de producción de cada nota para incrementar la cantidad de notas totales). De aquí se deriva directamente que necesita contar con un flujo de ingresos dado para poder solventar este tipo de gastos. La posición respetada y consolidada que tiene en el mercado fue alcanzada con este tipo de noticias y las mismas se van a mantener. No decimos con esto que solo quiera ofrecer un tipo de noticias específica, por el contrario, buscan constantemente identificar los intereses cambiantes de los lectores para poder satisfacer la demanda de la mejor forma posible manteniendo siempre su perfil de calidad.

En el mundo pre-internet, el mercado publicitario se repartía entre diarios, televisión, vía pública y cine, principalmente. Con el devenir del internet, con su alcance casi universal, la publicidad pasó a manos de nuevos jugadores (aquellos que eran capaces de capturar, en un mismo medio, al mayor número de visitas por unidad de tiempo). Podríamos establecer un corte temporal alrededor de 2012. Es aquí donde las firmas deben tomar una decisión dicotómica: cobrar o no cobrar por el acceso.

El negocio impreso empieza a reducirse y estos lectores no migran necesariamente al sitio digital de la firma (existen otros jugadores buscando capturar esta audiencia ambivalente). Luego, es necesario compensar esta caída de ingresos. Deben crecer en el mercado en expansión, en el mercado digital. Es necesario ofrecer una propuesta de valor en un momento oportuno para lograr que la audiencia esté dispuesta a pagar por el acceso a noticias.

Con la publicidad **solo** no alcanza para hacer buen periodismo, nos han dicho en una de nuestras reuniones con referentes de la firma. Es por esto que casi de forma natural el acceso deja de ser gratuito en internet: si valoran las noticias es lógico que exista una disposición a pagar por las mismas. Este fenómeno no es propio de la región, mas por el contrario imita la dinámica del exterior, donde la pionera con esta política de negocios fue *The New York Times*.

Es la norma de la industria en el mundo este mercado mixto, sin embargo en economías como la nuestra está aun más exacerbado ya que los ingresos por publicidad relativo a su producto bruto son muy bajos. En mayo de 2021, los ingresos por publicidad eran del orden de los \$200 por cada 1000 notas leídas, mientras que el precio de suscripción fue aproximadamente de \$600. Estos valores absolutos no son inforativos en una economía como la nuestra, donde la oneda pierde valor adquisitivo mes a mes. No obstante, transmite la idea del gap vigente entre ambos tipos de posibles fuentes de ingresos.

Llegado a este punto de la descripción del mercado debemos mencionar los jugadores que compiten con la firma por los lectores. No hay que dejar de tener en cuenta que existe un jugador clave en esta estructura: Infobae. Esta firma tiene una gran porción del mercado y es de libre acceso, poniendo trabas implícitas al precio que puedan cobrar sus pares sin recaer en partes muy elásticas de la demanda.

Pasemos entonces al plano del estudio de mercado. El hecho que las principales dinámicas se desarrrollen vía dispositivos electrónicos trae aparejado una ventaja sin presedentes: la firma tiene información casi perfecta de los intereses de los lectores (a priori puede capturar todos sus movimientos realizados en cada sesión a la plataforma). Es decir, todo se puede medir y cuantificar, volviendo indispensable un análisis de la audiencia tanto en términos de rezagos como de expectativas.

En particular, en un primer nivel la plataforma en estudio cuenta con una unidad de medida denominada “Unique Browser (UB)” que consiste en cookies que identifican al usuario asociado a un dispositivo y navegador. Esto permite llevar registros individuales en el tiempo sobre cada uno de estos UB.

En un segundo nivel tenemos a los lectores que navegan logueados, por ejemplo desde una cuenta de Google. Esto permite realizar una identificación uno a uno independientemente del medio desde el cual acceda. Con ambos métodos es posible para la firma realizar un historial de consumo del usuario, pudiendo mostrarle noticias que “matcheen” mejor con su perfil.

Además, la firma recopila información sobre el medio por el cual accedió, ya sea Google Search, entrada directa por la pagina del medio, por redes sociales (redirigidos desde un tweet por ejemplo), por notificaciones al celular, entre otros. Este tipo de datos junto con navegaciones previas y notas leídas ayudarán a identificar y clasificar a los lectores.

Existe todo un sistema de análisis de los datos que excede nuestro estudio, sobre cuál es el modo más eficiente de usar esta información. Para no volver redundante esta etapa, nombramos solo a modo ilustrativo para pasar al estudio concreto del modelo de negocios.



### 3.1 Modelo de Negocios

Como ya fue mencionado en reiteradas ocasiones, la plataforma recibe ingresos por suscripción y por publicidad. Su objetivo es, por un lado, retener el mayor tiempo posible a los suscriptores como suscriptores. Por el otro, quiere maximizar la cantidad de notas leídas.

La firma establece lo que denominamos un *paywall* (variable de elección,  $\bar{x}$  en nuestro modelo teórico). Esta variable captura la cantidad de notas máximas que puede leer un Unique Browser en una ventana móvil de 28 días sin suscribirse al medio. Es importante enfatizar este concepto de ventana móvil. Cada vez que un usuario accede a la plataforma, es identificado por el sistema, el cual computa la cantidad de notas acumuladas leídas y determina si, al hacer el siguiente click, la nota estará disponible, o por el contrario recibirá un cartel informando que superó las notas liberadas, ofreciendo la suscripción a la plataforma.

Aquí es donde remarcamos que los lectores recurrentes saben de la existencia de estas restricciones operativas, aunque no es razonable pensar que tengan un control sobre ellas ya que no conocen la cantidad acumulada ni el *paywall* en cada momento del tiempo (información asimétrica). Nuestro modelo teórico desarrollado previamente sostiene que existe un *paywall* óptimo, que dada la dinámica del mercado, se asemeja al caso 3 del escenario de Second Best. No obstante, no hay que dejar de mencionar que en este tipo de negocios las decisiones no son siempre puramente económicas, sino también reaccionando en función al accionar de sus pares en el mercado, cuestiones de ética profesional y otro tipo de variables.

## 4 Análisis Empírico

Contamos con bases mensuales sobre todos los meses del 2019, donde para cada usuario, identificado con un ID, se indica, para el mes dado, variables según distintas categorías:

- i. Dimensión: mes; usuario ID; dispositivo desde el cual accede; país; ciudad; tipo de usuario (suscriptor, no suscriptor, no suscriptor sin restricción)
- ii. Visitas: total de visitas; separadas según momento del día (mañana-tarde-noche); según forma de acceso
- iii. Notas: notas leídas, clasificadas según tipo (opinión, política, espectáculos, deportes, otras)
- iv. Paywall: Cantidad de choques con el paywall en el mes
- v. Conversión: fecha, hora y tipo de nota que generó la conversión

No contamos con una gran variedad de covariables que nos permitan hacer predicciones ni inferencia causal. Por esta razón, nuestro análisis se va a centrar principalmente en una interpretación de estadísticos descriptivos y gráficos.

Vamos a explotar un experimento natural: de forma aleatoria se “liberó” a partir de septiembre de 2019 y hasta abril de 2020 a aproximadamente un 0.03% de los no suscriptores (clasificados como no suscriptores sin restricción). La firma buscó mantener el porcentaje de liberados y, cuando fuera posible, sobre los mismos ID para ver patrones temporales, es decir en septiembre 2019 se seleccionó un subsample a liberar, si en octubre ese mismo ID accedió a la plataforma tuvo nuevamente acceso irrestricto.

Esto nos va a permitir comparar, por un lado, cómo se modifica el volumen de lecturas sobre aquellos que optan por no pagar pero se les concede acceso ilimitado al contenido. En este sentido, podemos inferir sobre el nivel de ingresos que genera la firma por suscripción, en particular relativo a los ingresos marginales que podría facturar por publicidad si levantara el paywall sobre toda la masa de lectores.

Además, nos va a servir para testear el modelo teórico, nuestro objetivo será poder inferir un poco más sobre la distribución de los usuarios y los valores críticos. Donde consideremos que no contamos con suficientes datos, lo vamos a dejar indicado para abrir la posibilidad a futuras investigaciones. Un análisis clave será estudiar si las lecturas de los suscriptores (según nuestro modelo tienen un  $\theta > \theta_3$  que se corresponde con lecturas superiores a  $x^*(\theta_3)$ , siendo  $x(\theta)$  una función biyectiva<sup>12</sup>) difieren estadísticamente de los no suscriptores sin restricción (deberían tener un parámetro  $\theta < \theta_3$ , presentando un patrón de lecturas distinto - sujeto a posibles errores de medición-). Concretamente vamos a poder observar a aquellos que tienen un  $\theta \in [\theta_2\theta_3]$  que dado el paywall antes leían de forma subóptima.

Dado que la selección de estos usuarios fue realizada de forma aleatoria, la decisión de liberar a un usuario es ortogonal al  $\theta$  del mismo. Esta independencia nos permite hacer inferencia sobre

---

<sup>12</sup>Para esta tipología de consumidores existe una relación 1 a 1 entre la cantidad de notas leídas y el parámetro  $\theta$ .

qué porcentaje de la masa de no suscriptores modificaría su patrón de lectura. Es decir podríamos tener una aproximación de la masa de los potenciales suscriptores ( $F(\theta_3) - F(\theta_2)$ ).

Vamos a empezar por realizar un análisis general de estadísticos descriptivos de los datos, para ello nos concentramos en los meses donde se realizó el experimento para ser consistentes con las interpretaciones. En particular, cuando analicemos en planos mensuales vamos a utilizar el período de noviembre 2019, ya que, como vamos a mostrar, septiembre difiere de los otros meses, y diciembre es un mes muy atípico, por ser fin de año, vacaciones y contar con muchos feriados.

## 4.1 Dimensión

Sobre la primer categoría es relevante mencionar la proporción de cada tipo de usuario:

	sep	oct	nov	dic
no suscriptor	0.9834	0.9607	0.9622	0.9613
no_suscriptor_sin_restricción	0.0095	0.0325	0.0313	0.0322
suscriptor	0.0071	0.0069	0.0065	0.0065

Table 1: Proporción de tipos de usuario. Período 2019

Podemos notar que las proporciones no sufren grandes sobresaltos en el último trimestre del año. Si existieran meses donde aumentara el número de suscriptores sería razonable que, de igual manera, se incrementen los lectores. En cuanto a los liberados, la proporción es una variable de decisión de la firma (se justifica el cambio de septiembre a octubre).

## 4.2 Visitas

Sobre la segunda categoría, exponemos la proporción según el tipo, y la volatilidad, midiendo media, desvío y mediana.

tipo	var	n	share	mean	std_error	median
no_suscriptor	n_sessions	1310691	0.962	2.229	4.714	1.000
no_suscriptor_sin_restriccion	n_sessions	42621	0.031	2.773	6.499	1.000
suscriptor	n_sessions	8805	0.006	21.648	28.741	10.000

Table 2: Estadísticos descriptivos. Número de sesiones según tipo de usuario. Período Nov-2019

Además, podemos estudiar el comportamiento de cada grupo a lo largo del día:

tipo	n	var	mean	std_error	std_error_norm	median
no_suscriptor	1310691	mañana	0.180	0.345	1.924	0.000
no_susc_sin_rest	42621	mañana	0.198	0.351	1.775	0.000
suscriptor	8805	mañana	0.251	0.263	1.049	0.200
no_suscriptor	1310691	tarde	0.369	0.436	1.182	0.000
no_susc_sin_rest	42621	tarde	0.368	0.425	1.157	0.091
suscriptor	8805	tarde	0.399	0.291	0.729	0.368
no_suscriptor	1310691	noche	0.452	0.451	1.000	0.364
no_susc_sin_rest	42621	noche	0.435	0.439	1.009	0.333
suscriptor	8805	noche	0.350	0.294	0.841	0.333

Table 3: Proporción de sesiones a los largo del día según tipo de usuario. Nov-2019

En este punto es interesante notar como difieren en comportamiento los distintos tipos de usuarios. La tabla muestra cómo se distribuyen en el día las sesiones (visitas a la plataforma) para cada tipo. Por ejemplo, el 18.4% de los no suscriptores leen por la mañana, el 36.9% por la tarde y el 45.2% restante por la noche. La media de los liberados y los restringidos son muy similares. Si miramos los desvíos estándar todos los grupos son muy volátiles, pero los suscriptores registran menores movimientos. En este punto, consideramos que el desvío normalizado a unidades de la media nos permite realizar una comparación más fehaciente de los datos, para evitar sesgos por los niveles propios de cada variable/tipo de usuario. Si observamos la sexta columna reafirmamos que los suscriptores son menos volátiles alrededor de su media. En todos los casos la mediana es menor que la media, lo cual es muy razonable ya que las distribuciones tienen cola a derecha.

En este punto es razonable preguntarse cómo se distribuyen la cantidad de sesiones en la muestra. Con este objetivo vamos a graficar las curvas de distribución acumulada de los tres grupos en un mismo gráfico. Naturalmente, los no suscriptores van a acumular mucha masa a izquierda ya que a partir del paywall no podrán seguir leyendo. No obstante, podrían ingresar a la página para chocarse con el mismo.

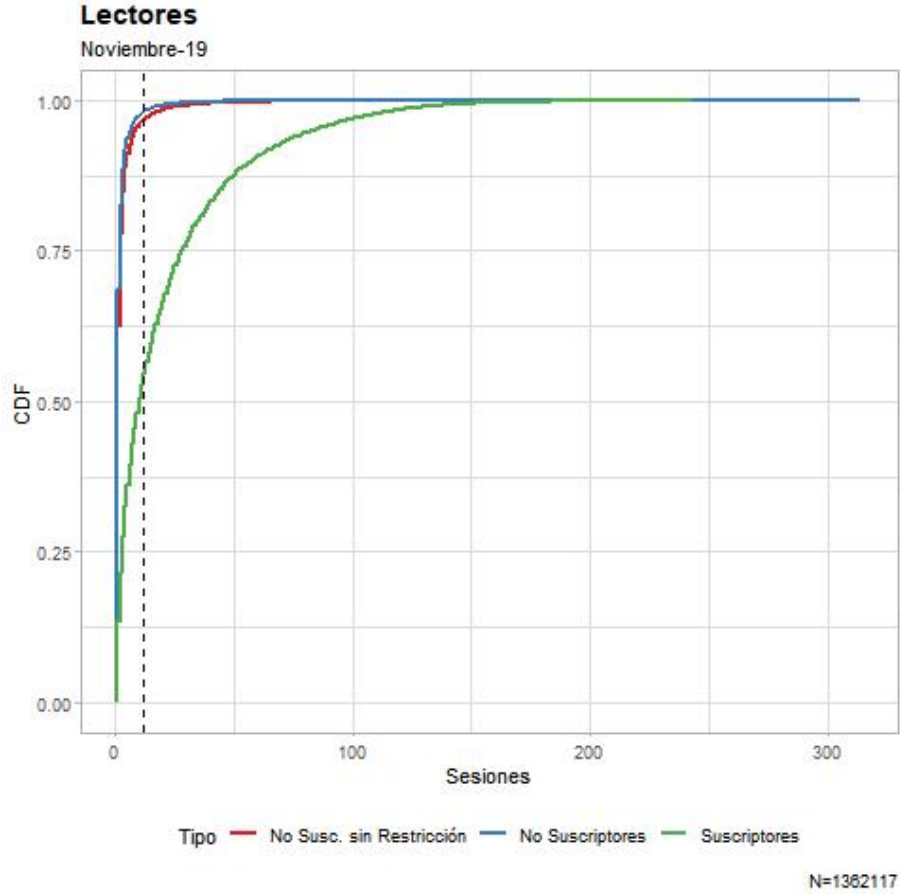


Figure 5: Función de distribución acumulada para cantidad de sesiones  
*Se indica en línea punteada el paywall a modo ilustrativo, ya que el mismo está medido en notas leídas las cuales pueden diferir de las sesiones.*

Por último, nos interesa testear si el comportamiento de los no suscriptores se modifica al asignarles un acceso irrestricto. Es decir, queremos testear si las CDF de estos grupos son estadísticamente distintas. Con este objetivo vamos a realizar el test de Kolmogorov-Smirnov, donde la hipótesis nula implica igual distribución.

Como se puede ver en la tabla (4) tenemos evidencia suficiente para rechazar la hipótesis nula, lo que nos llevaría a concluir que el paywall afecta el comportamiento de los no suscriptores. En particular, como podemos apreciar en el gráfico, la curva roja está por debajo de la azul, esto implica que los suscriptores no restringidos realizan más sesiones que sus pares restringidos. En conclusión el paywall es operativo sobre los usuarios y limita su actividad en la plataforma.

	statistic	p.value	alternative	method
KS test	0.06	0.00	two-sided	Two-sample Kolmogorov-Smirnov test

Table 4: KS test: Sesiones de no suscriptores vs no susc. sin restricción. Nov-2019

Un detalle no menor es la diferencia entre estos usuarios liberados y los suscriptores. Cabe

recordar que los usuarios no conocen su condición, razón por la cual en equilibrio podrían condicionar su comportamiento (por ejemplo por experiencias previas ante choques con el Paywall). Con esto queremos remarcar que nuestro modelo es estático, cuando en la realidad las decisiones son intertemporales. Para poder realizar una comparación real de estos dos grupos, los no suscriptores sin restricción deberían incorporar su condición a la hora de decidir la cantidad de notas que consumen por unidad de tiempo. Los suscriptores conoce su condición, ya que pagan un *fee* para identificarse como tales, para que ambos grupos sean comparables solo deberían distar en su decisión dicotómica de suscribirse/no suscribirse a la plataforma.

No contamos con suficientes períodos sobre los cuales un subsample de no suscriptores sean liberados de forma sostenida de manera tal que puedan actualizar sus creencias incorporando que pueden leer ilimitadamente. Por esta razón, consideramos que un análisis estático era la mejor forma de modelar nuestra idea y testearla en los datos.

### 4.3 Notas

En tercer lugar se encuentra la categoría más relevante para nuestro análisis: la cantidad de notas leídas. En este plano vamos a poder distinguir la variable agregada de las subcategorías internas. En particular, sería relevante saber si los suscriptores leen en promedio categorías de notas distintas de aquellos que no pagan un *fee*. Si este fuera el caso, la firma debería explotar esta información, realizando políticas tendientes a asociar a estos usuarios a la plataforma.

Veremos a continuación los estadísticos descriptivos y en segundo lugar los gráficos. Aquí recordamos que, siguiendo nuestro modelo, existe un biyección entre  $\theta$  y  $x^*(\theta)$  lo que nos ayudaría a intuir dónde podrían ubicarse los valores críticos.

tipo	var	n	share	mean	std_error	std_error_norm	median
no_suscriptor	n_notas	1325537	0.983	2.015	4.053	2.012	1
no_susc_sin_rest	n_notas	12825	0.010	3.935	6.866	1.745	2
suscriptor	n_notas	9613	0.007	35.684	61.793	1.732	13

Table 5: Estadísticos descriptivos. Número de notas leídas según tipo de usuario. Período Nov-2019

En primer lugar, notar que la media de los no suscriptores dista del paywall (12 en nuestra base de datos) significativamente. En particular, el hecho que la mediana se ubique en 1 y 2 nos indica que las lecturas no se están acumulando alrededor del límite. Esto se condice con nuestro modelo si pensamos que una masa significativa de los usuarios están leyendo su cantidad óptima (restricción no operativa).

En cuanto a los liberados, podemos ver que son más volátiles, aunque en niveles su variable también es mayor. Por esta razón, computamos el desvío estándar normalizado para tener una noción no sesgada por las unidades. Si observamos la septima columna concluimos que en realidad los restringidos son más volátiles. Este punto refuerza nuestra idea: aquellos que óptimamente resolvían leyendo por debajo del paywall lo siguen haciendo, luego la masa a izquierda no debería sufrir grandes cambios. No obstante, aquellos que se chocaban con el Paywall ahora pueden seguir

leyendo, permitiendo observaciones mayores. Esto, naturalmente, se verá reflejado tanto en la media como en el desvío. Pero cuando de forma conjunta se incorporan los dos estadísticos, el aumento de la media permite corregir el desvío bajando el nivel.

Ahora veamos como se distribuyen los tipos de usuarios con cada tipo de nota. La firma clasifica las notas como: política, opinión, deportes, espectáculos y otros. En la tabla 6 exponemos una descripción de las mismas:

tipo	n	var	mean	std_error	std_error_norm	median
no_suscriptor	1310691	política	0.063	0.212	3.337	0.000
no_susc_sin_rest	42621	política	0.070	0.214	3.042	0.000
suscriptor	8805	política	0.158	0.194	1.230	0.111
no_suscriptor	1310691	opinión	0.012	0.097	8.054	0.000
no_susc_sin_rest	42621	opinión	0.012	0.093	7.720	0.000
suscriptor	8805	opinión	0.039	0.111	2.844	0.000
no_suscriptor	1310691	deportes	0.123	0.296	2.413	0.000
no_susc_sin_rest	42621	deportes	0.131	0.296	2.252	0.000
suscriptor	8805	deportes	0.133	0.185	1.385	0.071
no_suscriptor	1310691	espectáculos	0.208	0.364	1.752	0.000
no_susc_sin_rest	42621	espectáculos	0.247	0.378	1.529	0.000
suscriptor	8805	espectáculos	0.159	0.190	1.199	0.118
no_suscriptor	1310691	otro	0.475	0.456	0.960	0.500
no_susc_sin_rest	42621	otro	0.494	0.441	0.892	0.500
suscriptor	8805	otro	0.463	0.278	0.600	0.472

Table 6: Estadísticos Descriptivos. Tipo de notas según tipo de usuario. Período Nov-2019

La plataforma con la que estamos trabajando es una referente en notas de política y opinión. Podemos observar que las media para los suscriptores se diferencia de la de los no suscriptores en esta categoría y lo contrario sucede sobre las notas de espectáculos. Por otro lado, si miramos la voltilidad de los usuarios, los suscriptores son más consistentes con sus lecturas ya que, para cada categoría, sus desvíos estándar normalizados son menores.

A su vez, los valores para los no suscriptores son muy similares entre ambos tipos. Esto nos sugiere que los tipos de usuarios no difieren únicamente en el volumen leído, como predice nuestro modelo, sino también en el tipo de consumo que realizan. Aquí existen, en principio, dos posibles explicaciones: o bien los usuarios que más leen tienen este tipo de preferencias, es decir estamos viendo una correlación y no causalidad, o bien este tipo de categorías se distinguen de las otras, digamos en calidad, y esto genera que aquellos que presentan preferencias sobre estos tipos resuelven leyendo más. En este plano estaríamos queriendo imponer la causalidad como: política/opinión  $\rightarrow$  calidad  $\rightarrow$  más lecturas  $\rightarrow$  Paywall operativo  $\rightarrow$  suscripción.

El hecho que prácticamente la mitad de la notas leídas estén categorizadas como “otros” nos limita en gran medida el análisis en este plano.

Recapitulando nuestros resultados teóricos, argumentamos que aquellos que no se suscriben es porque la cantidad óptima que resuelve su problema es lo suficientemente chica, donde suficientemente chica quiere decir que la utilidad marginal de seguir leyendo por encima del paywall es menor al precio de suscripción. Luego, estos usuarios tienen un  $\theta < \theta_3$ .

Identificar en la muestra a  $\theta_2$  no es un problema ya que se determina implícitamente por el paywall (son aquellos usuarios que están chocando la línea punteada). El verdadero desafío está en poder afirmar algo sobre  $\theta_3$  ya que no hay, a priori, ninguna señal clara del mismo<sup>13</sup>. Pero siguiendo nuestro modelo teórico, los no suscriptores deberían acumular prácticamente toda su masa (dejando lugar a errores de medición) antes de este valor.

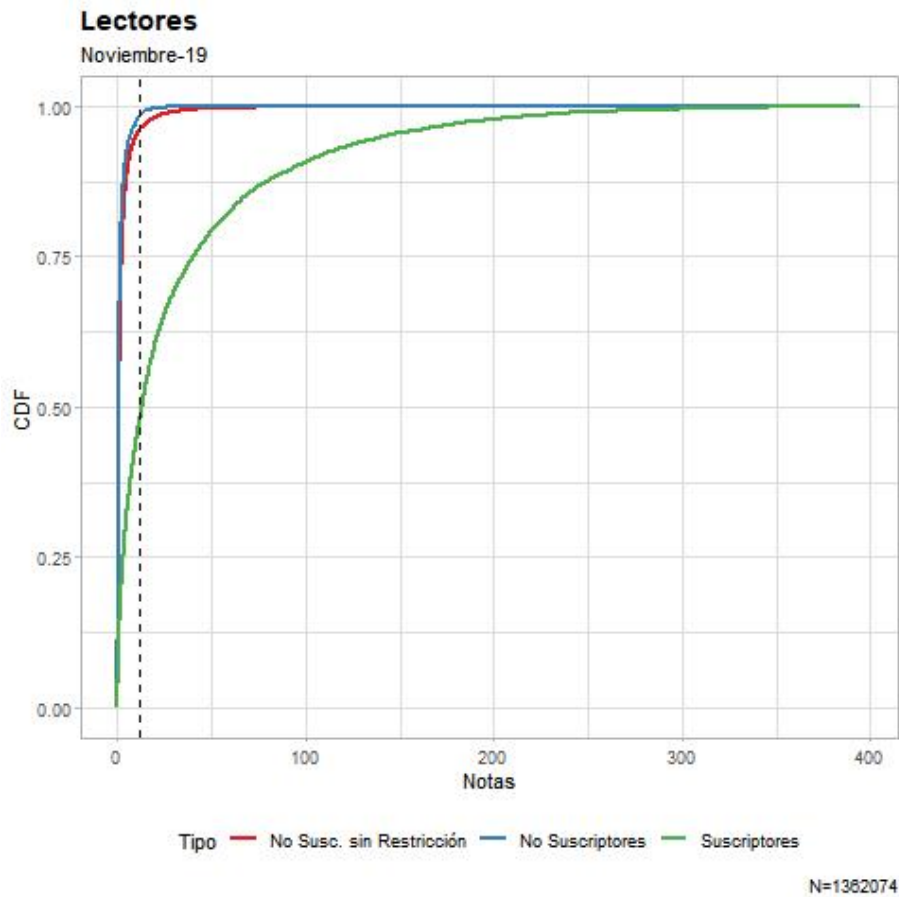


Figure 6: Función de distribución acumulada para cantidad de notas

*Se indica en línea punteada el paywall, la cantidad de observaciones está indicada en el gráfico. Las mismas corresponden al período de noviembre 2019. Para mantener la escala del gráfico excluimos del mismo las observaciones que se correspondían con más de 400 lecturas en un mes (43 sobre una base de 1362117).*

En la Figura 6 graficamos de forma conjunta los 3 tipos de usuario. Podemos ver que liberar no

<sup>13</sup>En este punto, estamos queriendo identificarlos dada la política actual de la firma, que impone el Paywall en 12 notas y el precio de suscripción en \$600 mensuales.



suscriptores (curva roja) corre la distribución a la derecha pero igualmente a partir de un cierto nivel parecerían converger ambas curvas al 1 (acumulan toda su masa). Con esta imagen, reafirmamos lo mencionado relativo a las sesiones, el libre acceso no genera que la curva de distribución se asemeje con los suscriptores, dando una señal de que existen características en los usuarios que los vuelven de un tipo o del otro ( $\theta$  en nuestro modelo teórico, un análisis más exhaustivo nos llevaría a aumentar la dimensión de  $\theta$ , donde un vector permitiría capturar características distintas en distintas dimensiones)<sup>14</sup>.

Lo que querríamos medir es la efectividad del paywall para restringir las lecturas y generar incentivos a la suscripción. Por esta razón, nos interesa nuevamente testear si ambas distribuciones son estadísticamente distintas. Computamos el test de Kolmogorov-Smirnov y nuevamente rechazamos la hipótesis nula ya que el p-value es suficientemente chico.

	statistic	p.value	alternative	method
KS_test	0.10	0.00	two-sided	Two-sample Kolmogorov-Smirnov test

Table 7: KS test: Notas de no suscriptores vs no susc. sin restricción. Nov-2019

Concluimos de nuestro análisis que existe un grupo de no suscriptores que modifican su comportamiento si les dejan leer de forma irrestricta. El estadístico que computa el test de Kolmogorov-Smirnov consiste en computar el supremo entre ambas distribuciones. Sea  $G(x)$  la función de distribución acumulada de la cantidad de notas leídas de los no suscriptores con Paywall operativo y  $H(x)$  su análogo sobre los liberados. Luego, el valor crítico será:

$$KS = \sup_x |G(x) - H(x)|$$

Es decir, para cada nivel de notas computa la distancia máxima entre ambas curvas. Naturalmente, dado que estamos mirando funciones de distribución acumulada, el estadístico toma valores entre  $[0,1]$ , siendo 0 cuando ambas distribuciones son la misma, y 1, en el otro extremo, cuando la distancia es máxima.

Sobre la cantidad de notas leídas el estadístico es 0.10 y sobre las sesiones 0.06. Concluimos que sobre el total de notas leídas el comportamiento reacciona más. Este resultado es razonable ya que para una misma sesión esperamos múltiples notas leídas.

Luego, la política de la firma está resignando ingresos por publicidad sobre estas potenciales notas leídas en un trade-off con los ingresos por suscripción.

Desde un punto de vista puramente matemático podemos computar este valor. Vamos a estimar este volumen de potenciales ingresos por publicidad para tener una idea más tangible del problema que enfrenta la firma en el trade-off de paywall y precio de suscripción.

<sup>14</sup>Si tuvieramos más características de los usuarios se podrían inferir con mayor seguridad. Por ejemplo, si tuvieramos algún proxy al nivel socioeconómico, sobre los suscriptores la tarjeta con la cual se suscriben es una buena primera aproximación ya que las mismas están categorizadas sobre el nivel de gasto que realiza el cliente con el banco de referencia. Sobre los no suscriptores necesitaríamos un análisis paralelo para poder compararlos en este plano.

Dados nuestros estadísticos descriptivos, sabemos que, en promedio, un no suscriptor sujeto al Paywall lee 2.015 notas en el mes. Además, si de forma aleatoria ( $\theta$  igualmente distribuidos) liberamos un subsample de esta población, la media pasa a ubicarse en 3.935. Es decir, imponer la restricción genera que cada no suscriptor lea en promedio 1.92 notas menos en el mes. Luego, la firma está resignando  $1.92 \cdot q \cdot \%_{NS}$  por mes de ingresos por publicidad a cambio de generar  $p \cdot \%_S$  ingresos por suscripción, siendo  $\%_{NS}$  el porcentaje de no suscriptores de la plataforma, y  $\%_S$  el porcentaje de suscriptores.

Si la firma eligió esta política debería ser cierto que el primer monto es menor al segundo. Sabemos que la suscripción es aproximadamente 600 por mes, y que el espacio publicitario paga 200 por cada 1000 notas leídas, es decir  $q = 0.2$ . Luego computamos:

$$\begin{aligned} & \text{Ing Publicidad vs Ing Suscripción + Publicidad} \\ & q \cdot (mean^{lib} - mean^{rest}) \cdot \%_{no.suscriptores} \quad vs \quad p \cdot \%_{suscriptores} \\ & 0.2 \cdot 1.92 \cdot 0.993 \quad vs \quad 600 \cdot 0.007 \\ & 0.3813 \quad vs \quad 4.2 \end{aligned}$$

Nuevamente resaltamos que esta es una primera aproximación, dado que los usuarios no conocen su condición de liberados, es decir estamos pensando que el gap de los valores medios podría ser mayor si la información no fuera asimétrica entre ambas partes del mercado. Este análisis fue debidamente desarrollado en el modelo teórico, de forma gráfica y analítica para poder interpretar de forma visual los montos computados (casos 1 y 3). No volveremos a comentarlo para no ser redundantes con este punto. La aproximación numérica apoya nuestra teoría.

Además, podríamos computar hasta que nivel de gap, dado los nivel de  $q$  y  $p$  sería óptimo para la firma elegir el esquema Paywall + Suscripción (caso 3) por sobre Lecturas irrestrictas (caso 1)<sup>15</sup>:

$$\begin{aligned} & \text{Ing Publicidad} < \text{Ing Suscripción + Publicidad} \\ & q \cdot (mean^{lib} - mean^{rest}) \cdot \%_{no.suscriptores} < p \cdot \%_{suscriptores} \\ & 0.2 \cdot (mean^{lib} - mean^{rest}) \cdot 0.993 < 600 \cdot 0.007 \\ & (mean^{lib} - mean^{rest}) < \frac{600 \cdot 0.007}{0.2 \cdot 0.993} \\ & (mean^{lib} - mean^{rest}) < 21.14 \end{aligned}$$

No obstante, todas las variables están interrelacionadas, con lo cual esta desigualdad es una primera aproximación al problema de la firma.

#### 4.4 Paywall

Por último, podemos realizar un análisis de los choques con el Paywall que sufren los no suscriptores (único tipo de usuario sujeto a esta restricción).

En la base de datos contamos con 3 features relativos al Paywall:

---

<sup>15</sup>Mantenido a  $\bar{x} = 12$ ,  $p = \$600$  y  $q = \$0.2$ , donde los primeros 2 valores son variables de elección de la firma.

- Paywall Metered: El paywall se presentó al intentar leer una nota adicional
- Paywall Metered Entrada: El paywall se presentó al usuario al ingresar a la plataforma (antes de elegir la nota a leer)
- Paywall Voluntario: Voluntariamente tocaron el botón de suscribirse

Al estudiar estas covariables llegamos a resultados inconsistentes, razón por la cual no podemos confiar en la validez de estos registros.

Para el mes en análisis, el paywall implica que en una ventana móvil de 28 días un usuario no suscripto puede leer a lo sumo 12 notas. Nuestra base es mensual, correspondiente en este caso a noviembre 2019. Este mes tiene 30 días. En este sentido, existe margen para que se “renueven” las notas a leer por los dos días adicionales. Si el usuario agotara sus notas en los primeros dos días (el caso más extremo) podría leer a lo sumo 24 notas.

Con esto en mente nuestro análisis fue el siguiente:

- En una primera instancia, sobre las tres variables estudiamos que porcentaje de los no suscriptores chocaban el paywall.
- En segundo lugar, realizamos chequeos de consistencia. Calculamos los porcentajes para la submuestra de aquellos usuarios que, durante todo el mes de noviembre leyeron más de 12 notas. Los resultados no fueron consistentes, ya que solo el 58.4% estaba chocando con el Paywall.
- Dados los resultados obtenidos, extendimos el análisis a submuestras cada vez más restrictivas: para más de 18 y más de 24 notas.

	Paywall_NS	notas12	notas18	notas24
N	1310691	20205	6001	3111
Paywall Metered	0.0253	0.5840	0.5244	0.4719
Paywall Metered Entrada	0.0149	0.3186	0.2611	0.2372
Paywall Voluntario	0.0005	0.0035	0.0045	0.0039

Table 8: Promedio de choques al paywall. Período Nov-2019

Los resultados obtenidos son inconsistentes. Sobre la última columna, el total de los usuarios deberían haber enfrentado el paywall, no obstante los registros indican que a menos de la mitad de esta población le fue mostrado. El tamaño de esta muestra es chico pero al retroceder en las columnas, es decir al ser menos restrictivo en el recorte de la muestra, pero por encima del paywall, el porcentaje no disminuye.

A partir de estos resultados no podemos confiar en esta categoría, ya que sugieren errores de medición sostenidos, al menos sobre aquellos que leen muchas notas, siendo estos grupos sobre los que tenemos particular interés.

## 4.5 Aproximación

La firma no puede tomar decisiones eficientes si no conoce la distribución de los usuarios. Por esta razón, levanta base de datos para intentar estudiar de la mejor manera la dinámica del mercado. En esta instancia vamos a proponernos inferir dónde se ubica  $\theta_3$  (lector indiferente entre suscribirse o no a la plataforma).

Debemos realizar una serie de supuestos de identificación para poder computar la posición de este agente crítico. Dada la información con la que contamos y el modelo que desarrollamos vamos a suponer que:

- Existen errores de medición en los datos. Suponemos que los mismos son ortogonales a la cantidad de notas leídas y al tipo de usuario. El sistema que impone la restricción de lecturas no tiene porqué fallar, por esta razón la hipótesis más razonable es que la fuente de errores se genera al consolidar los datos levantados. Luego, no existe en principio razones para suponer que este tipo de errores correlacione de forma distinta con cada tipo de usuario.
- El 100% de los no suscriptores deberían acumularse a izquierda del Paywall. Luego, el porcentaje que lo supere debe estar explicado por un error de medición. Vamos a computar este error  $\epsilon$  dado por la proporción de no suscriptores que lean por encima de la restricción impuesta por la firma.

Con esto en mente, vamos a graficar la Figura 6 en el sector de interés para poder explicar nuestra forma de pensar la dinámica antes de computar.

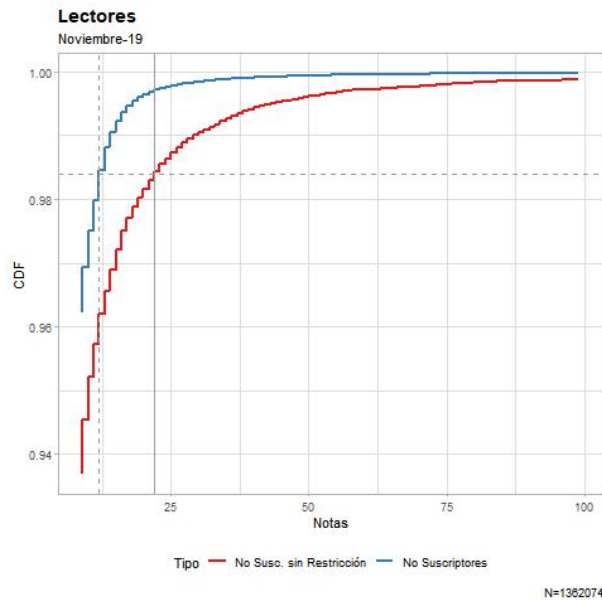


Figure 7: No suscriptores. Período Nov-2019

En la Figura 7 graficamos la CDF de los dos tipos de no suscriptores. La línea punteada vertical identifica el Paywall. El cruce de esta curva con la CDF de los No Suscriptores restringidos (curva

azul) nos da un proxy del error de medición ( $\epsilon$ ) de la base de datos. Luego, con la línea punteada horizontal identificamos hasta que nivel de distribución acumulada vamos a considerar para buscar el  $\theta_3$ .

Por último, la línea vertical ubicada en 22 notas identifica para qué nivel de notas se acumula  $(1 - \epsilon)\%$  de la distribución de los no suscriptores liberados. Con esta estrategia de identificación llegaríamos al resultado de que  $\hat{\theta}_3$ , dado  $\bar{x} = 12$  y  $p = \$600$ , es aquel agente que resuelve para 22 notas óptimas.

Este método es naturalmente criticable, ya que estamos realizando supuestos de identificación no testeables en la práctica. Para agregar robustez, vamos a replicar el análisis para los meses de octubre y diciembre, para descartar que exista una gran volatilidad en el resultado según el período estudiado.

En la figura (8) graficamos las CDF correspondientes al período previo y posterior. La primer observación a realizar es que el error  $\epsilon$  de los tres meses se mantiene. La curva horizontal está graficada en los tres casos en la misma ubicación (0.984). Por otro lado, vemos que la ubicación de  $x(\hat{\theta}_3)$  se desplaza a izquierda o derecha del 22 en cada caso. En particular, para octubre se acumula el  $(1 - \epsilon)\%$  de la distribución de los no suscriptores liberados en 21 notas, mientras que para diciembre se corre 3 lugares a la derecha, ubicándose en 25.

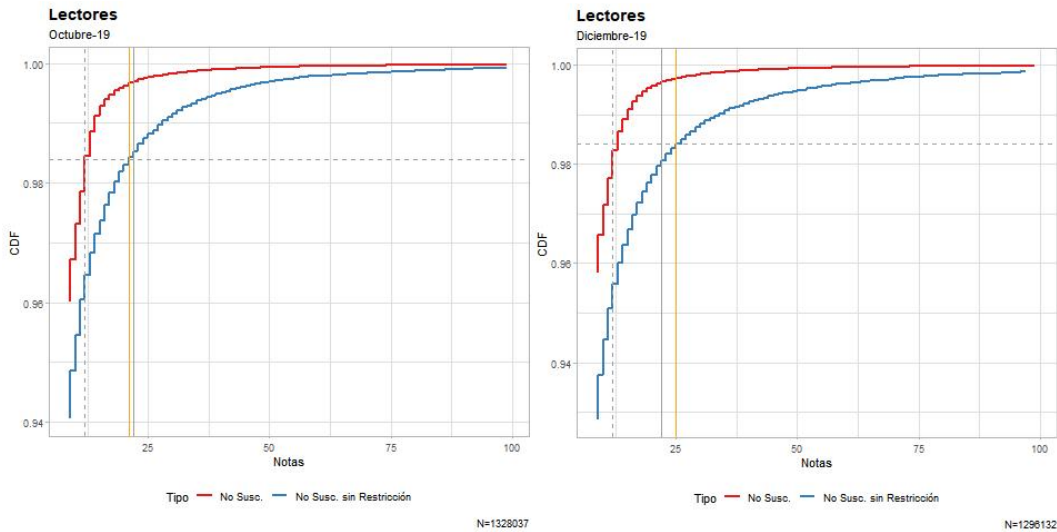


Figure 8: No suscriptores.

*El gráfico de la izquierda corresponde al período de octubre 2019 mientras que el de la derecha a diciembre 2019. Las curvas horizontales y verticales en gris se ubican en los mismos lugares que las correspondientes al gráfico de noviembre 2019.*

En conclusión, con esta primer aproximación concluimos que, con un 98.4% de confianza, los no suscriptores son aquellos lectores que resuelven leyendo a lo sumo el doble del paywall. Que resuelvan óptimamente de esta manera justifica que no quieran incurrir en un costo fijo de suscripción.

## 5 Suscriptores

En la sección previa nos propusimos realizar un resumen descriptivo de la base de datos sobre todos los lectores de la plataforma. En esta sección nos vamos a concentrar puntualmente en los suscriptores. La política de la firma es cobrar un *fee* a este grupo ya que, como mostramos numéricamente, los ingresos son mayores con esta estrategia de negocios frente a una posible alternativa de levantar las restricciones de acceso y maximizar los ingresos por publicidad.

Los suscriptores son uno de los agentes claves en nuestro modelo. Por esta razón, nos interesa tener la mayor noción de este grupo en torno a nuestra variable de interés: cantidad de notas leídas. Nuestro modelo teórico sostiene que se suscriben aquellos que su cantidad óptima de lecturas es lo suficientemente grande, donde suficientemente grande implica, en primer lugar que la cantidad de notas que resuelve su problema es mayor al paywall y además mayor a  $x(\theta_3)$ , es decir una cantidad tal que la utilidad marginal de leer por encima del paywall es mayor al *fee* de suscripción.

No obstante, como nombramos en la sección previa, nuestro modelo es estático, cuando en la realidad los suscriptores interactúan con la plataforma período a período. Por esta razón, en el siguiente se indica para cada mes del 2019, qué porcentaje de suscriptores lee una cantidad menor al paywall (12 notas):

	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Share	0.44	0.45	0.41	0.42	0.43	0.43	0.44	0.45	0.45	0.44	0.44	0.44

Table 9: Porcentaje de Suscriptores por debajo del Paywall

Frente a esta situación vemos que nuestro modelo no *matchea*, en principio, con la base de datos. Por esta razón, vamos a enunciar dos posibles hipótesis (compatibles entre si):

- Los suscriptores leen en valor esperado por encima del paywall, razón por la cual se justifica la suscripción a la plataforma. Sin embargo, si vemos el comportamiento mes a mes y de forma agregada esto nos podría sesgar sobre la dinámica detrás de los datos. Para testear esta hipótesis querríamos buscar estacionalidad en el comportamiento de los lectores, tanto a nivel individual como agregado. Sería necesario contar con datos para años previos para poder agregar robustez a los resultados (poder desligar el efecto de la PASO en agosto 2019 del comportamiento promedio propio de este mes, por ejemplo).
- La plataforma presenta fuertes switching cost, generando que los suscriptores paguen el *fee* por no incurrir en el costo de tener que desvincularse de la misma. Para testear esta hipótesis necesitaríamos ver a los suscriptores en el tiempo y esperaríamos una caída sostenida de lecturas en el tiempo, por una serie de meses estarían sostenidamente leyendo notas menores al paywall para finalmente cancelar la suscripción.

En este punto, es válido realizar un análisis econométrico de nuestra base de datos. Nuestra estrategia va a consistir en realizar regresiones sobre los cuantiles y sobre la media, para un modelo de datos de panel, con efectos fijos por individuo (primeras diferencias).

La regresión cuantílica es razonable ya que la media no siempre es una medida representativa de la población. Este método consiste en minimizar las desviaciones absolutas ponderadas con pesos asimétricos. En particular, el algoritmo para el cuantil  $\theta$  resuelve el siguiente problema:

$$\min_{\beta_\theta} \sum_{y_i \geq \beta_\theta X_i} \theta |Y_i - \beta_\theta X_i| + \sum_{y_i < \beta_\theta X_i} (1 - \theta) |Y_i - \beta_\theta X_i| \quad (21)$$

La tabla 10 expone los coeficientes estimados en cada caso, siendo  $t_j$  dummies correspondientes a cada mes del año. La constante corresponde al estimador de enero, y cada coeficiente se interpreta relativo a este mes, es decir un  $\hat{\beta}_j > 0$  implica que ese mes, en promedio, se leyeron más cantidad de notas que en enero (si miramos la regresión sobre la media, en los otros casos el análisis será sobre los cuantiles). Notar que los efectos de nivel de los individuos los neteamos al tomar primeras diferencias para evitar resultados sesgados por esta causa.

En particular, nos interesa la significatividad de los coeficientes y su cambio a medida que nos movemos de cuantil. A partir de este análisis, podemos observar que los cuantiles más altos son muy consistentes con su nivel de lectura en el tiempo, en particular para el caso del cuantil 95, en ningún mes tiene lugar un cambio significativo en cantidad de notas leídas relativo a enero. Este grupo de lectores es muy fiel a la plataforma. Los desvíos estándar de este grupo son grandes pero debemos interpretarlos relativo a la media (niveles).

Mirando el otro extremo, el 5% con menores lecturas, presentan los cambios en niveles más amplios relativo a enero y en prácticamente todos los meses significativos con un nivel de confianza del 1%, dando una gran noción de la volatilidad de este grupo. Es interesante notar que la constante de este grupo es -83.25 mientras que en el otro extremo es 69.75 (un diferencia de 153 puntos).

Por otro lado, dejando por fuera los cuantil extremos, pareciera existir una estacionalidad en torno a los meses de verano, donde la cantidad de notas son relativamente pocas y no presentan grandes cambios. Este comportamiento parece razonable, dado que esperamos una mayor cantidad de gente de vacaciones, menos acontecimientos importante en la agenda (ya sea político, deportivo, etc), entre otras causas. A partir de marzo, las notas leídas empiezan aumentar de forma estadísticamente significativa en todos los casos, donde el pico, para el 2019, se alcanza entre marzo y abril y empieza a caer de forma gradual. En este punto, dado que solo contamos con datos para un año, no podemos asegurar si este comportamiento es propio del mes o si se debe a algún shock externo.

¿Cómo explicamos que en marzo y abril, meses que no aparentan ser particularmente distintos relativo al resto del año, se registraron las mayores lecturas del año? Es probable que el comienzo del año académico y la vuelta a la actividad reactiven la lecturas, como características propias de estos meses entérminos generales. Hechos puntuales de 2019, podemos nombrar en abril el incendio que marcará la historia de París: la catedral de Notre Dame. Este acontecimiento podría ser un desestabilizador en las lecturas, generando que este mes se distinga de marzo y mayo.

Sin embargo, dado que nuestra firma se encuentra dentro de las referentes en el plano político argentino, es interesante notar que en mayo del 2019 no aumentaron de forma significativa relativo a abril las notas leídas, siendo que Cristina Fernández de Kirchner anunció su candidatura a mediados

	Media	q05	q25	q50	q75	q95
(Intercept)	-6.41*** (1.27)	-83.25*** (5.62)	-24.75*** (1.86)	-4.50*** (0.80)	11.17*** (1.76)	69.75*** (10.87)
t2	-2.73 (1.79)	9.17 (6.84)	0.33 (2.64)	-2.08* (1.18)	-4.67** (2.25)	-15.33 (12.78)
t3	18.28*** (1.79)	44.92*** (6.20)	17.50*** (2.20)	11.42*** (1.44)	16.92*** (2.77)	16.92 (12.49)
t4	18.22*** (1.79)	44.92*** (6.42)	16.75*** (2.09)	9.33*** (1.37)	16.50*** (2.63)	16.17 (13.32)
t5	11.04*** (1.79)	42.67*** (6.30)	14.50*** (2.04)	4.25*** (1.02)	3.92 (2.43)	-5.92 (12.71)
t6	10.11*** (1.79)	45.67*** (6.98)	14.50*** (2.02)	4.08*** (0.99)	3.33 (2.38)	-13.00 (11.40)
t7	8.78*** (1.79)	44.92*** (6.16)	13.58*** (2.05)	3.58*** (0.94)	1.67 (2.24)	-11.83 (11.70)
t8	3.47* (1.79)	37.33*** (5.91)	10.83*** (2.07)	1.67* (0.99)	-2.75 (2.05)	-29.58*** (11.16)
t9	6.26*** (1.79)	40.00*** (6.45)	11.75*** (2.06)	2.42** (0.95)	1.58 (2.18)	-21.33* (12.00)
t10	6.37*** (1.79)	36.83*** (7.14)	10.67*** (2.16)	1.92* (0.99)	0.58 (2.14)	-9.33 (11.54)
t11	2.13 (1.79)	30.42*** (6.21)	7.50*** (2.15)	0.58 (1.02)	-2.25 (2.10)	-26.33** (11.50)
t12	-5.25*** (1.80)	20.92*** (7.77)	0.58 (2.19)	-3.17*** (1.12)	-9.00*** (1.95)	-37.00*** (12.04)
$R^2$	0.04					
Adj. $R^2$	0.04					
Num. obs.	10128	10128	10128	10128	10128	10128
Percentile		0.05	0.25	0.50	0.75	0.95

\*\*\* $p < 0.01$ ; \*\* $p < 0.05$ ; \* $p < 0.1$

Table 10: Regresiones Efectos Fijos. Período 2019. Tamaño de muestra N=10128

de este mes. A su vez, en agosto el plano económico del país se desplomó como consecuencia de los resultados de las PASO, nuevamente si miramos relativo al mes anterior no encontramos aumentos. A partir de este mes, podríamos esperar un aumento en las lecturas, relativo a otros años dada la gran incertidumbre que se instauró en el país sobre el futuro color político que ocuparía el Poder Ejecutivo.

En octubre, estallidos sociales se desataron en Chile. Con este acontecimiento, naturalmente, gran parte de la sociedad argentina se movilizó dadas las violaciones a los Derechos Humanos que se estaban viviendo en nuestro país vecino. La movilización y represión militar fueron una gran



preocupación para la región.

Finalmente, en diciembre esperamos que se crucen los efectos de las vacaciones con el cambio de presidente y la primeras notas que hablarían de un nuevo virus que estaba matando de forma incontrolable a la población china.

Para cerrar esta idea, veamos gráficamente la evolución de los coeficientes estimados:

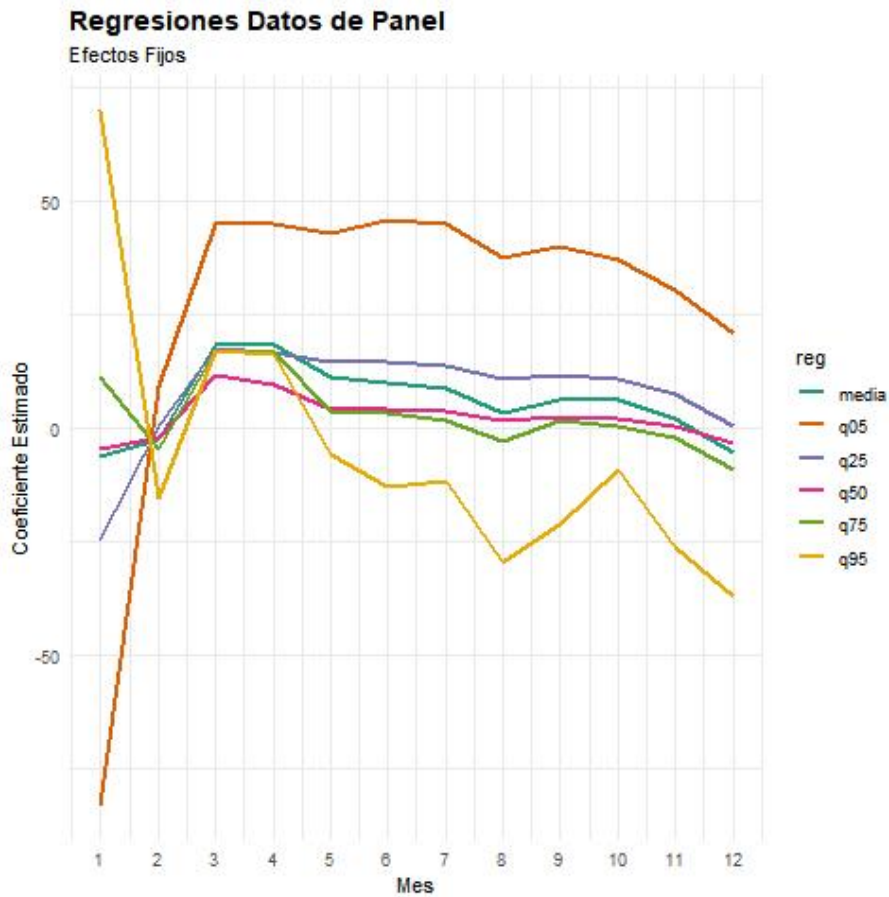


Figure 9: Coeficientes mensuales estimados. Período 2019

*Las estimaciones fueron realizadas mes a mes, realizamos un gráfico continuo para poder seguir la tendencia en cada caso.*

Por otro lado, para intentar testear nuestra segunda hipótesis querríamos seguir a los mismos individuos en el tiempo. Con este fin, identificamos en la muestra a aquellos usuarios para los cuales tenemos observaciones en todos los meses. Si miramos a nivel individual, la figura 10 nos muestra la heterogeneidad de los lectores en el tiempo, donde ya no parece tan obvio el aumento en marzo/abril.

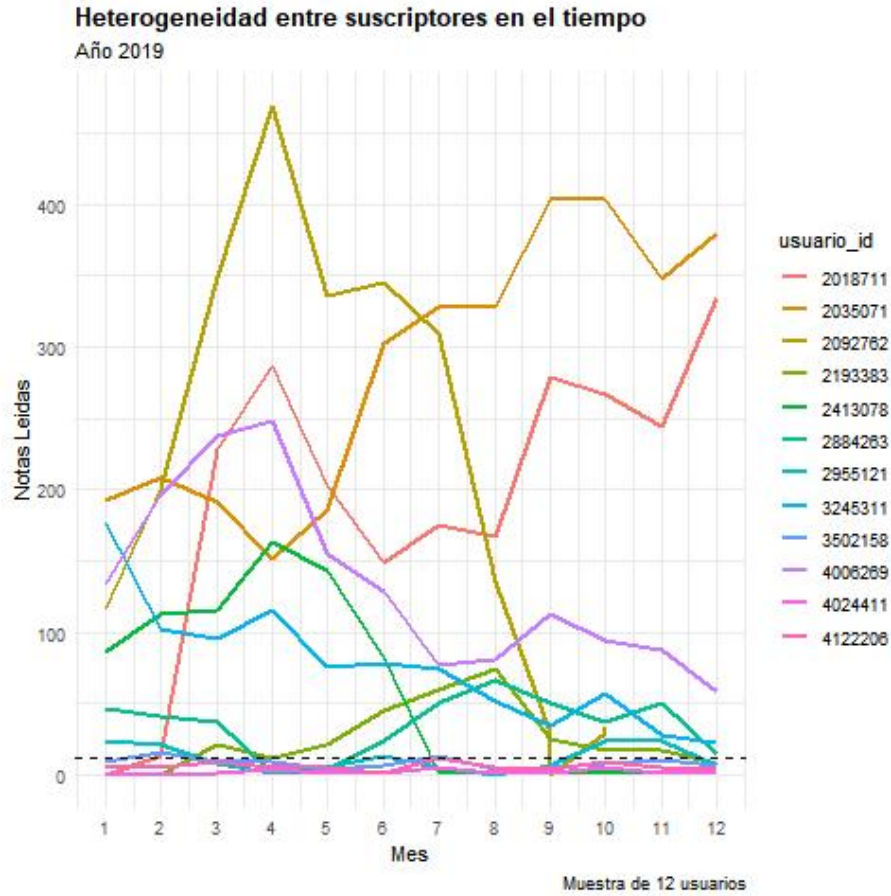


Figure 10: Muestra de 12 suscriptores

*El gráfico muestra, para 12 suscriptores elegidos de forma aleatoria, como cambian sus lecturas a través del 2019. Realizamos la selección de individuos de forma aleatoria y cambiando los grupos el patrón en los datos se mantenía, la imagen es meramente ilustrativa.*

A su vez, esperamos que la relación con la plataforma genere un nivel de *engagement* sobre los usuarios. Este punto puede ser controversial. Por un lado, mayor tiempo accediendo al diario genera que lea más, porque desarrollo interés. A su vez, en un punto me canso y disminuyen las lecturas (según nuestra hipótesis de *switchig cost*). Con este fin vamos a exponer una comparación entre el grupo de usuarios que observo durante todos los meses (podemos tener seguridad que llevan en promedio 6 meses suscriptos) *vs* la muestra entera (donde tenemos muchos usuarios que observamos en menos oportunidades).

En particular vamos a exponer un gráfico de caja (Figura 11) que nos da una noción de la volatilidad en las lecturas, para el 50% de la muestra de cada grupo (ancho de cada caja). Aquí podemos ver nuevamente el patrón alrededor de marzo/abril. Además, la imagen agrega robustez a la teoría de que el *engagement* genera consistencia (aumento sostenido de las lecturas). Podemos ver que en todos los casos, como enunciaba nuestra primer hipótesis, en promedio leen por encima del *paywall*.

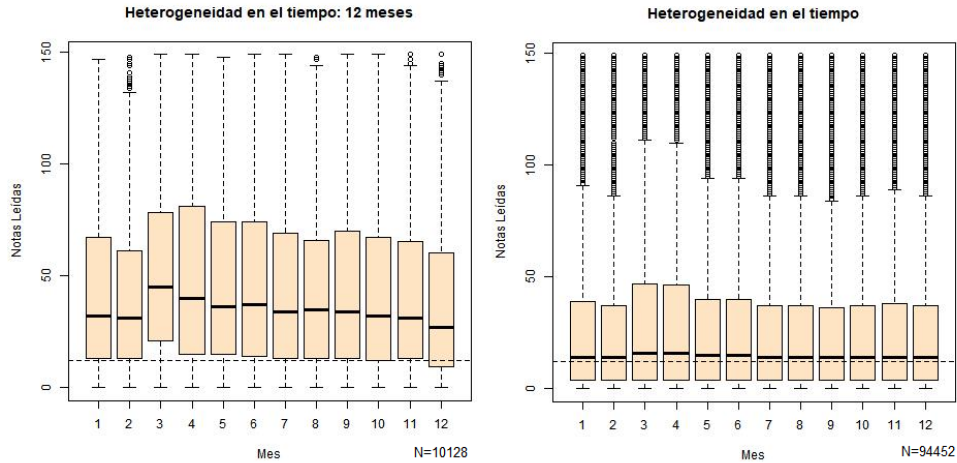


Figure 11: Gráfico de Caja

*El gráfico de la izquierda corresponde a aquellos usuarios que se observan durante todo el año, mientras que el de la derecha en para el total de suscriptores. En línea punteada se representa el Paywall. Para mantener la escala del gráfico quitamos de la imagen los bordes, aunque existen outliers en ambos grupos.*

Por último, vamos a exponer un gráfico del estilo de control sintético (Figura 12), donde nuestros sintéticos serán los cuantiles de notas en cada momentos. Vamos a utilizar los mismos niveles (5, 25, 50, 75, 95). Estas curvas están graficados en naranja mientras que en azul se muestra la media. El eje Y mide la cantidad de notas de cada suscriptor (sin netear de la media). Podemos notar que en este escenario solo el cuantil 05 queda por debajo del paywall (línea punteada).

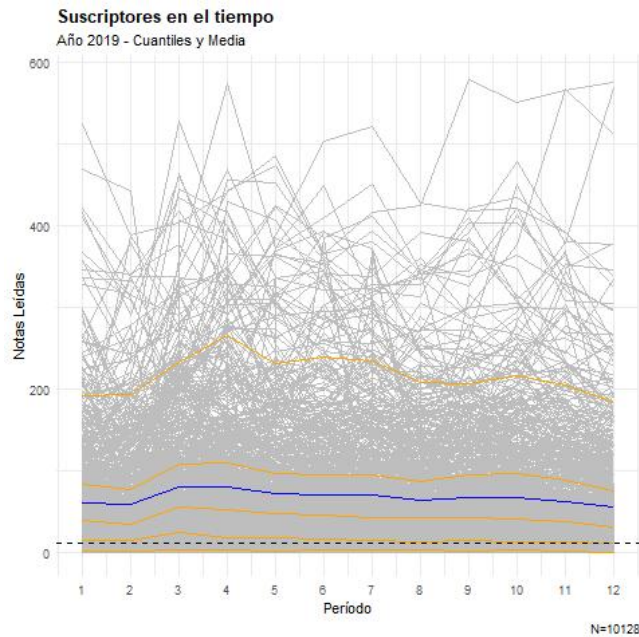


Figure 12: Suscriptores en el tiempo

## 6 Demanda

El análisis hasta este punto se basó en nuestras creencias sobre cómo se podría modelar la utilidad de los lectores para construir la demanda que enfrentaba la firma para cada par  $(\bar{x}, p)$  ofrecido en el mercado. Una vez construido nuestro modelo, buscamos testear los resultados de forma empírica. No obstante, existen múltiples razones que podrían explicar los datos, y nuestra hipótesis es una de ellas.

¿Cómo agregamos robustez a nuestros resultados? Lo ideal sería analizar el otro lado del mercado. Estudiar a los consumidores con mayor precisión, testear hipótesis de cómo toman las decisiones y qué factores tendrían mayor poder explicativo para un modelo dado. En este plano, deberíamos realizar encuestas para formarnos el complemento a la base de datos que tenemos. Esta opción no está disponible a nuestro, razón por la cual vamos a abrir la discusión a futuras investigaciones y buscar en la literatura trabajos que hayan investigado fenómenos similares.

Desde el punto de vista de la firma, podría hacer una mayor inversión en recolectar datos más informativos de los usuarios<sup>16</sup>. No obstante, esto no sería suficiente ya que estaríamos enfrentando un problema de endogeneidad análogo al clásico problema de querer explicar salarios y nivel de empleo. Al observar a los usuarios, estamos enfrentando una demanda que decidió consumir en esta plataforma para cada tipo de política de la firma y contenido ofrecido. Con esto queremos enfatizar que sería una visión sesgada del mercado.

Un problema adicional en este plano es el ritmo de evolución que tiene este mercado. Artículos que investigaron este fenómeno hace 10 años están obsoletos, ya que no incorporan el rol del *smartphone* en la dinámica. Además, el propio vínculo de los usuarios con las plataformas está en constante cambio. La disposición a pagar por el servicio (*Willingness To Pay*, *WTP*) ha mutado y nada nos sugiere que ya se encuentre en un punto estacionario. Los niveles de suscripción a plataformas como Netflix o Spotify están en constante crecimiento. Una década atrás este fenómeno hubiese sido impensado. Estos cambios de comportamiento en los patrones de consumo de los agentes van a repercutir en la relación a establecer con las plataformas de noticias. Fletcher & Nielsen (2020) testean precisamente este problema y concluyen que existe una asociación positiva a nivel individual entre pagar por otro tipo de contenido en línea (películas, música, libros) con pagar por noticias en línea. Adicionalmente, encuentran una asociación positiva entre pagar específicamente por entretenimiento en línea y pagar por las noticias.

¿Por qué existe resistencia a la suscripción (paga) por un servicio que valoro? La respuesta a esta pregunta está ligada al rubro a analizar. Puntualmente relativo a las noticias, la literatura pone el mayor foco en la presencia de sustitutos cercanos de libre acceso. Acá se abre entonces la siguiente discusión:

¿Es la firma capaz de generar una diferenciación vertical (desde el punto de vista del usuario) sobre el producto ofrecido relativo a los otros sitios de internet? Los consumidores racionales no están dispuestos a pagar por algo a lo que podrían acceder gratis, y noticias en la red se obtienen a un

---

<sup>16</sup>Es de nuestro conocimiento que la firma cuenta con features adicionales. Por un tema de confidencial trabajamos en el corriente paper con aquellos que pudieron proveernos.

costo de búsqueda nulo. Luego, si las noticias de cada plataforma son sustitutos lo suficientemente cercanos, ¿por qué alguien estaría dispuesto a suscribirse? Además, la elasticidad precio cruzado es particularmente alta en estos mercados (los precios de suscripción están prácticamente acotados).

Por otro lado, desde el surgimiento del internet se ha desarrollado en paralelo la “cultura de lo gratis” (Miguel de Bustos 2010). Este fenómeno acompañado de una idea colectiva de falta de valor a las noticias por su creación barata, no diferenciable, con periodistas contando las mismas historias de la mismas fuentes converge en equilibrio a tasas muy bajas de suscripción. Adicionalmente, las redes sociales están desplazando a las plataformas de noticias como fuente de información. Desde este nuevo espacio se consolida la información, se abren debates y se lleva seguimiento de la agenda pública.

De forma complementaria, Zhang & Nguyen (2004) agregan a los causantes los métodos de pago disponibles. Este punto, aunque surge de un paper relativamente antiguo, mas no por eso es un resultado obsoleto. Existe, particularmente en sociedades como la nuestra donde no siempre están garantizados los niveles de confidencialidad necesarios, resistencia a completar datos personales y de tarjetas de crédito o débito en páginas de internet.

En paralelo al fenómeno mencionado, tiene lugar el hecho que los medios generan ingresos por publicidad (el lector no incorpora que debe ser él quien pague por el servicio dado que está indirectamente generando ingresos por lectura vía publicidad). No obstante, esta fuente se vio reducida sustancialmente en los últimos años. Como ya nombramos previamente, por la capacidad de otras plataformas de reunir a más usuarios en un mismo lugar, y en un plano no menor por el paso de la computadora al celular. Las pantallas se redujeron drásticamente en tamaño, quitando posibles espacios publicitarios (Berger, Steininger & Hess 2015).

Por último, antes de pasar a resultados de la literatura, cabe aclarar que no hay que subestimar el fenómeno cultural (los papers en general trabajan sobre bases de datos americanas). Con este objetivo en mente, con los resultados de “Reuters Institute Digital News Report 2020 (RIDNR 2020)” veamos los principales indicadores para Argentina<sup>17</sup>:

Relativo a Estados Unidos y España se reporta el resultado a la siguiente pregunta:

*Which, if any, of the following have you used in the last week as a source of news?*

	Online (incl. social media)	TV	Social Media	Radio	Print
USA	73%	60%	47%	21%	26%
Spain	83%	71%	63%	24%	28%
Argentina	90%	77%	78%	24%	30%

Table 11: Fuente: RIDNR 2020. Base abril 2020

Para leer los datos hay que tener en cuenta que en abril 2020 el cierre en Argentina era total, junto con la incertidumbre del manejo de la pandemia, razón que podría explicar una mayor búsqueda de información.

<sup>17</sup>Este reporte recolectó datos de áreas urbanas, teniendo validez externa sobre este sector de la población.

Manteniendo los países del control, ante la pregunta:

*Have you paid for ONLINE news content, or accessed a paid-for ONLINE news service in the last year?* los resultados fueron:

USA	Spain	Argentina
20%	12%	11%

Table 12: Fuente: RIDNR 2020

Sugiriendo un WTP menor en nuestra sociedad.

Por último, veamos indicadores propios de Argentina:

Relativo a fuentes digitales, buscaron en la siguientes fuentes según la proporción indicada:

Infobae	40%
TN online	36%
Clarín online	29%
La Nación online	22%
Minuto Uno	15%

Table 13: Weekly reach online. Fuente: RIDNR 2020

La presencia de Infobae fue discutida previamente, consolida un gran porcentaje de la audiencia, principalmente por el hecho de ser de libre acceso. El mismo fenómeno explicaría las búsquedas en TN Online. Le siguen los portales pagos tradicionales, donde Clarín se distingue significativamente de La Nación en 7 puntos porcentuales.

No podemos distinguir de estos datos las intersecciones para saber qué porcentaje de los entrevistados está utilizando estas plataformas para informarse. No obstante, nos da una buena primera aproximación.

Por último, el resultado del reporte relativo a Argentina sostiene:

*Changing Media: Social media have overtaken TV news amongst our online sample for the first time, while weekly print consumption has fallen from 45% to just 23% over the last three years. Eight out of ten (80%) respondents say they now use the smartphone to access news each week.*

Hechos los paralelismos, veamos cuáles son los determinantes del WTP.

Chyi (2012) desarrolla un modelo de OLS para explicar la intención de pago sobre la versión impresa, la versión web y en formato de apps de las noticias. Las principales covariables utilizadas son: *Gender (female), age, education, income, news interest, print newspaper use, online news use, TV news use.*

En ninguna de las 3 regresiones las variables demográficas resultan significativas. Por otro lado, el género y la edad tienen influencia en los modelos de la versión web y las apps. En particular, gender es significativa a niveles de confianza del 99% y 95% respectivamente en cada modelo con signo negativo. Es decir, los hombres tienen mayor WTP. En cuanto a la edad, razonablemente

también se predice un estimador negativo, ya que los jóvenes son los más dispuestos a pagar por este nuevo servicio. El nivel educativo y los ingresos tienen coeficientes positivos aunque en estos modelos no resultan significativos con los niveles de confianza tradicionales (99%, 95%, 90%).

La variable que mide el interés por noticias resulta significativa y positiva (razonablemente). Por último, nuevamente no resulta sorprendente que aquellos que se identifican con consumo de noticias online arrojan coeficientes positivos y significativo para los formatos de web y apps, mientras que su contraparte en la versión impresa es negativa y significativa con un nivel de confianza del 99%.

En línea con estos resultados, Goyones (2014) corre una regresión logística donde por variable resultado se utiliza la dummy que vale 1 si se está dispuesto a pagar por noticias en línea. Los controles son análogos a los mencionados en el modelo anterior. De igual manera, no encuentran evidencia de un efecto por las variables demográficas. Los efectos sobre la edad son negativos, y en este caso el género no tiene poder explicativo. Ese fenómeno lo explican, por un lado porque los jóvenes están más conectados que los viejos. Por otro, la penetración de *smartphones* y tablets es mucho mayor sobre este grupo etario. Por último, encuentran que mayores niveles de ingreso tienen un efecto positivo en el WTP.

En una segunda escala, el trabajo mencionado busca medir si el uso de las redes sociales (medido con Twitter) prediciría la intención de pago. Concluyen que las personas que usan Twitter de manera rutinaria (todos los días) tienen dos a tres veces más probabilidades de visitar un sitio web de noticias que la persona promedio (esto no significa que estén más dispuestos a pagar por noticias en línea).

En conclusión, nos queda pendiente estudiar el otro lado del mercado. Existen múltiples trabajos que abordan esta pregunta de investigación y los resultados son razonables. Los jóvenes se desarrollaron en paralelo con esta revolución tecnológica, tomando por natural estos nuevos fenómenos. En consecuencia, sus niveles de WTP son mayores y esto los ubica entre los más dispuestos a pagar por noticias en línea.

Además, el hecho que existan otras plataformas exitosas en consolidar a la audiencia, por ejemplo de música o películas/series ofreciendo entretenimiento virtual pago, está generando que un mayor porcentaje de la sociedad recurra a estos recursos, naturalizándolos. La bibliografía consultada nos indica que este cambio de comportamiento va a repercutir de igual manera en el WTP de las noticias.

De esta manera, concluimos que existe mucho espacio para investigar y que una firma que enfrenta esta disyuntiva debe considerar que este mercado está en constante cambio. Para sacar el mayor rédito posible deberá ser capaz de adaptarse rápidamente a las demandas cambiantes, ofrecer un servicio de calidad, distintivo y atractivo para inducir al lector a tomar la decisión dicotómica de suscribirse.

## 7 Conclusiones

El presente trabajo tuvo por objeto estudiar y caracterizar la estrategia de negocios óptima para un diario digital en el marco del *trade-off* entre ingresos publicitarios e ingresos por suscripción. Para ello utilizamos tanto un enfoque teórico como uno empírico.

Con respecto al modelo teórico incursionamos en la literatura de mercados multilaterales y de modelos de discriminación por calidad. Analizamos la decisión de una plataforma, en un contexto de asimetría informativa, que cobra un *fee* por suscripción y libera una cierta cantidad de artículos a un continuo de consumidores heterogéneos en una única dimensión: su valoración por artículos. En particular, analizamos tres estrategias posibles:

- La primera consiste en liberar todos los artículos, de modo tal de maximizar el número de lecturas, lograr masividad, y financiarse únicamente con ingresos por publicidad. En este escenario, la decisión de la plataforma maximiza el excedente agregado y no existe pérdida de eficiencia.
- Una estrategia alternativa consiste en cobrar un precio de suscripción y solo permitir la lectura si un consumidor decide suscribirse. Esta estrategia atenta contra la masividad y además genera pérdida de eficiencia, ya que consumidores con una valoración positiva del bien no pueden leer en equilibrio.
- Por último, analizamos una estrategia en la cual existe un límite de artículos que se pueden leer sin cargo, pero para leer una cantidad mayor de notas debe suscribirse a la plataforma. Esta política comercial permite alcanzar una masividad mayor que la del punto anterior, capitalizar mayores ingresos publicitarios, y si bien permite un mayor número de lecturas en equilibrio, existen consumidores que leen una cantidad de artículos ineficientemente baja. Este caso comprende un escenario de *versioning*, en el cual la firma ofrece un producto de calidad alta (suscripción) y una versión dañada del mismo (artículos disponibles sin cargo), y con esta estrategia permite que los consumidores se auto-clasifiquen en grupos según sus disposición a pagar (que comprende una característica fundamental para la firma, sin embargo, es no observable).

La plataforma utilizará la estrategia que maximice sus beneficios y eso dependerá de una gran cantidad de variables, en particular de la distribución de valoraciones por parte de los consumidores y del ingreso que reciba por nota leída en concepto de publicidad.

Por otro lado, el estudio empírico nos permitió convalidar ciertos resultados del modelo teórico:

- Los suscriptores difieren sistemáticamente de los no suscriptores en el nivel de notas que resuelven para leer en equilibrio. Aún en un escenario liberado, con un nivel de confianza del orden del 98% leen menos de 25 notas por mes.
- Difieren también en el tipo de consumo que realizan. En particular, en nuestra base de datos encontramos que los suscriptores son los que consumen en mayor proporción notas de política y opinión. Queda pendiente determinar si existe una causalidad que los relacione.



- Numéricamente, y dada la evidencia que vemos en la realidad, para una firma como la analizada le es óptimo ubicarse en un escenario donde exista un paywall positivo y un *fee* por suscripción (en buena medida por los ingresos generados por la lectura y por la publicidad).
- Los jóvenes serían los más dispuestos a pagar por contenido en línea. Resultados de trabajos consultados sugieren, razonablemente, que el mercado digital está en constante evolución. Por esta razón, las firmas que operen en el largo plazo serán aquellas capaces de adaptarse a estos cambios. Resolverán ofreciendo un paquete atractivo para que los usuarios se auto-clasifiquen y de esta manera, podrán sacar el mayor rédito posible al mercado.

A modo de cierre, podemos sostener que el mercado de artículos digitales es un mercado complejo, que involucra dinámicas de permanente cambio y los jugadores deben adaptar sus estrategias constantemente. La digitalización es un fenómeno que avanza rápido en el mundo, y el que se queda atrás, pierde.

## 8 Apéndice

En este apéndice se encuentran los desarrollos matemáticos de los resultados de la sección 2.

### Proposición 2

*La masa de suscriptores disminuye con  $\bar{x}$  y disminuye con  $p$ .*

La masa de suscriptores es  $1 - F(\theta_3)$ ; luego la misma es decreciente en  $\theta_3$ . A partir de la ecuación (4) definimos

$$H(\theta_3, \bar{x}, p) = \theta_3 \phi(x^*(\theta_3)) - p - k(x^*(\theta_3)) - \theta_3 \phi(\bar{x}) + k(\bar{x}) = 0$$

Para analizar cómo depende  $\theta_3$  de la cantidad de artículos sin cargo  $\bar{x}$ , aplicamos el Teorema de la Función Implícita

$$\begin{aligned} \frac{\partial \theta_3}{\partial \bar{x}} &= - \frac{k'(\bar{x}) - \theta_3 \phi'(\bar{x})}{\phi(x^*(\theta_3)) + \theta_3 x^{*'}(\theta_3) [\phi'(x^*(\theta_3)) - k'(x^*(\theta_3))] - \phi(\bar{x})} \\ &= \frac{\theta_3 \phi'(\bar{x}) - k'(\bar{x})}{\phi(x^*(\theta_3)) - \phi(\bar{x}) + \theta_3 x^{*'}(\theta_3) \underbrace{[\phi'(x^*(\theta_3)) - k'(x^*(\theta_3))]}_{=0 \text{ por ecuación (1)}}} \\ &= \frac{\theta_3 \phi'(\bar{x}) - k'(\bar{x})}{\underbrace{\phi(x^*(\theta_3)) - \phi(\bar{x})}_{>0 \text{ porque } \phi(\cdot) \text{ es creciente}}} \end{aligned}$$

Para analizar el signo del numerador, definimos  $g(x; \theta_3) = \theta_3 \phi'(x) - k'(x)$ . Por la ecuación (1) sabemos que  $g(x^*(\theta_3); \theta_3) = 0$  y como  $\phi(\cdot)$  es cóncava y  $k(\cdot)$  convexa podemos asegurar que  $g(\cdot)$  es decreciente. Luego, como  $g(x, \theta_3)$  es continua debe ser cierto que  $g(\bar{x}, \theta_3) > 0$ , por lo que el numerador es positivo y

$$\frac{\partial \theta_3}{\partial \bar{x}} > 0$$

Luego, dado que  $\theta_3$  es creciente en  $\bar{x}$ , la masa de suscriptores es decreciente en  $\bar{x}$ . Para analizar cómo depende  $\theta_3$  del costo de suscripción  $p$  utilizamos, nuevamente, el Teorema de la Función Implícita

$$\frac{\partial \theta_3}{\partial p} = - \frac{(-1)}{\underbrace{\phi(x^*(\theta_3)) - \phi(\bar{x})}_{>0 \text{ porque } \phi(\cdot) \text{ es creciente}}} > 0$$

Entonces, como  $\theta_3$  es creciente en  $p$ , la masa de suscriptores es decreciente en  $p$ .

## 9 Referencias

- Armstrong. 2006. "Competition in two-sided markets", RAND Journal of Economics, Vol. 37, No. 3, pp. 668-691
- Berger, Matt, Steininger & Hess. 2015. "It Is Not Just About Competition with "Free": Differences Between Content Formats in Consumer Preferences and Willingness to Pay". Journal of Management Information Systems, 32:3, 105-128
- Chyi. 2012. "Paying for What? How Much? And Why (Not)? Predictors of Paying Intent for Multiplatform Newspapers" The International Journal on Media Management, 14:227-250
- Deneckere & McAfee. 1996. "Damaged goods", Journal of Economics and Management Strategy 5(2): 149-174.
- Fletcher & Nielsen (2020). "Are Netflix and Spotify Subscribers More Likely to Pay for Online News? Comparative Analysis of Data From Six Countries" International Journal of Communication 14(2020), 3439-3457
- Goyanes. 2014. "An Empirical Study of Factors that Influence the Willingness to Pay for Online News" Journalism Practice, 2014 Vol. 8, No. 6, 742-757
- Küng, Newman & Picard. 2016. "Online news". In Handbook on the Economics of the Internet. Cheltenham, UK: Edward Elgar Publishing.
- McAfee. 2007. "Pricing Damaged Goods". Economics: The Open-Access, Open-Assessment E-Journal, 1 (2007-1): 1-19.
- Miguel de Bustos. 2010. "Sobre pago y gratuidad Momentos de crisis en la industria periodística" Revista Telos 83: 24-35
- Newman, Fletcher, Schulz, Andi & Kleis Nielsen. 2020. "Reuters Institute Digital News Report 2020". University of Oxford
- Tirole. 1988. "The Theory of Industrial Organization" MIT Press.
- Tirole & Rochet. 2003. "Platform competition in two-sided market", JEL: L5, L82, L86, L96.
- Varian. 1997- "Versioning Information Goods", in Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property, MIT Press, pp.190-202.
- Zhang & Nguyen. 2004. "What Makes Consumers Willing to Pay for Online Services? Discovering Consumer Opinions and Determinant Factors on Charged Online Services" Journal of Internet Commerce, 2:4, 35-53