

Please Subscribe to Continue Reading: An Empirical Analysis of Readers' Responses to Newspaper Paywalls

Eric Overby, Georgia Institute of Technology

Adi Pattabhiramaiah, Georgia Institute of Technology

Vamsi Kanuri, University of Notre Dame

Research Context

What happens when readers hit your paywall?

Do they:

1. Subscribe (or renew)
 2. Log into their pre-existing account
 3. Read non-premium articles
 4. Depart
-
5. Circumvent the paywall
 6. Probe the paywall
 7. Display obstinance

The Daily Tribune

Pavlov's Hot Dog Stand To Close

By: George Burdell, 11/11/2026

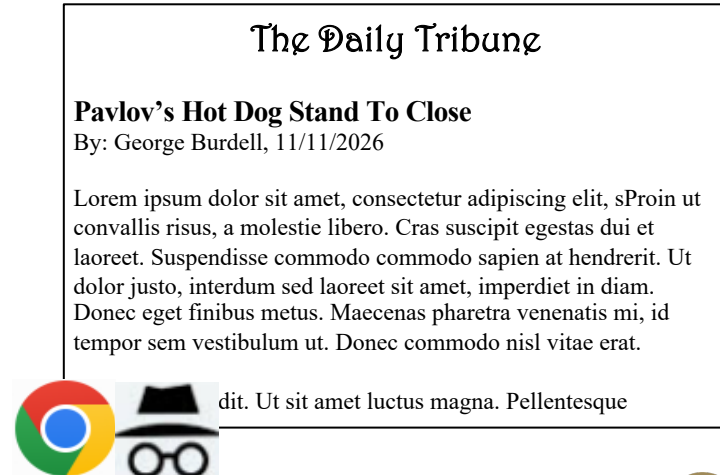
Subscribe to continue reading.
Already a subscriber? **Sign in**
Get every story for \$1 week

Subscribe

Why Is This Hard to Study?

Analyzing how readers respond to the paywall requires:

1. Detailed clickstream data, including when the paywall was presented.
2. A method to measure when a reader circumvents the paywall.
 - This is difficult to measure because it is a surreptitious behavior meant NOT to be observed.



Data and Variables

We use clickstream data from a top 30 North American newspaper, focusing on the period from April 1, 2021 to March 30, 2022.

209,170,500 article page views during this period.

- **8,934,456** (4.3%) of those triggered the stop page.
- These stop pages were triggered by **7,595,615** distinct browser cookies.

Data Structure

Date/ time	Cookie ID	IPv4	User ID	URL	Pre- mium?	De- vice	OS	Browser	Browser Version	Ref- errer	Event
5/12/21 13:44:03	59f3ad2 28a059f	50.238. 18.210	null	https://www.news.com/ nation/sports-team- wins/	1	PC	Wind ows	Chrome	85.0.413 .83	Direct	Paywall Stop

Measuring Responses in the Data

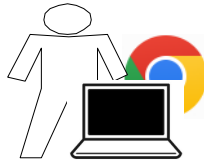
Reader responses are straightforward to observe, *except for circumvention*.

	Date/ time	Cookie ID	IPv4	User ID	URL	Premium	Event
Subscribe (or renew)	5/12/21 13:44:03	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ local/mayor-elected/	1	Stop
	5/12/21 13:44:31	59f3ad228a059f2b	50.238.18.210	22ake8	https://www.news.com/ local/mayor-elected/	1	Completed
Log into existing account	5/12/21 13:44:03	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ local/mayor-elected/	1	Stop
	5/12/21 13:44:31	59f3ad228a059f2b	50.238.18.210	58d6e7	https://www.news.com/ local/mayor-elected/	1	
Read non- premium articles	5/12/21 13:44:03	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ local/mayor-elected/	1	Stop
	5/12/21 13:44:31	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ new-covid-variant/	0	
Probe the paywall	5/12/21 13:44:03	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ local/mayor-elected/	1	Stop
	5/12/21 13:44:31	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ sports-team-wins/	1	Stop
Display obstinance	5/12/21 13:44:03	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ local/mayor-elected/	1	Stop
	5/12/21 13:44:31	59f3ad228a059f2b	50.238.18.210	null	https://www.news.com/ local/mayor-elected/	1	Stop

Measuring Circumvention: Client-Side Paywall

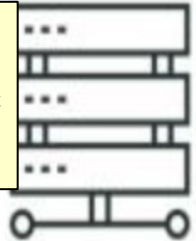
The newspaper we study had a **client-side** paywall (Marshall 2021).

1. Newspaper website serves entire article to the reader.



Pavlov's Hot Dog Stand To Close

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do



2. Newspaper website then checks reader's browser cookie to determine if article should be blocked via JavaScript.

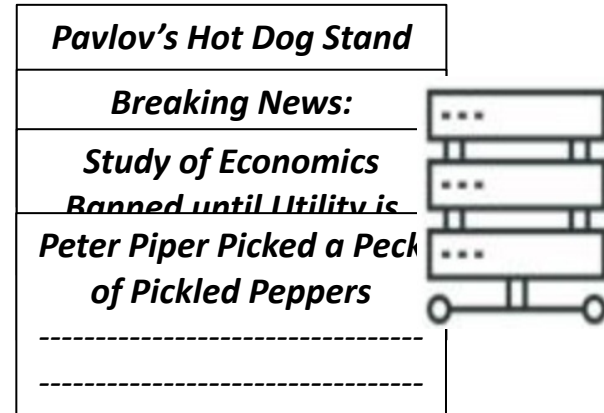


Measuring Circumvention: Metered Paywall

The newspaper we study had a **metered** paywall.

During our study period, the meter was either 2 or 3 free premium articles before the paywall was presented.

Meter = 3 articles



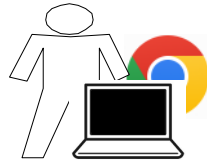
Measuring Circumvention: Main methods

Client-side, metered paywalls permit readers to circumvent the paywall to access the article that has been blocked. We focus on two main methods:

- 1) Disabling JavaScript.
- 2) Using a different browser cookie.

Circumvention By Disabling JavaScript

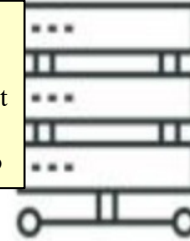
This type of circumvention exploits the design of client-side paywalls that serve the entire article content to the reader's browser only to then block it via JavaScript when relevant.



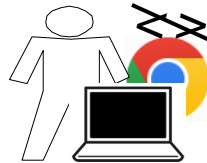
JavaScript determines whether to present the stop page

Pavlov's Hot Dog Stand To Close

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do

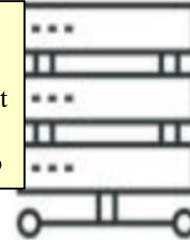


After reader disables JavaScript, the article is no longer obscured



Pavlov's Hot Dog Stand To Close

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do



Circumvention By Disabling JavaScript

This behavior appears in the data as shown below.

Date/ time	Cookie ID	IPv4	User ID	URL	Browser	Browser Version	Device	OS	Event
5/12/21 13:44:03	59f3ad22 8a059f2b	50.238. 18.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.4183. 83	PC	Windows	Stop
5/12/21 13:44:31	59f3ad22 8a059f2b	50.238. 18.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.4183. 83	PC	Windows	

Same cookie,
same IPv4,
same UserID

Accessing
same article

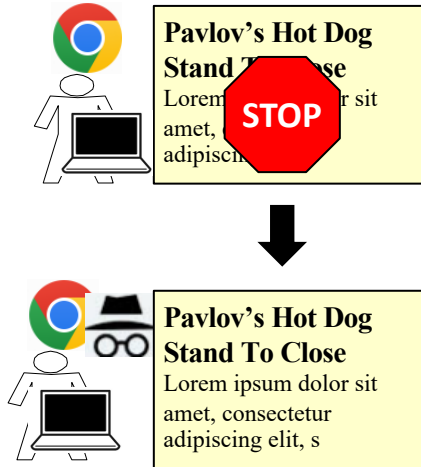
Same browser,
device, etc.

Appears to be same reader, but
second page request does NOT
generate another stop event.

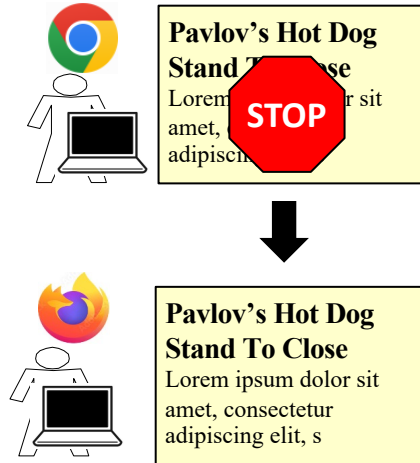
Circumvention By Using a Different Cookie

This type of circumvention exploits that paywalls use cookies to identify readers. Readers who get a different browser cookie present as a new reader, which resets the article “meter”, allows them to circumvent.

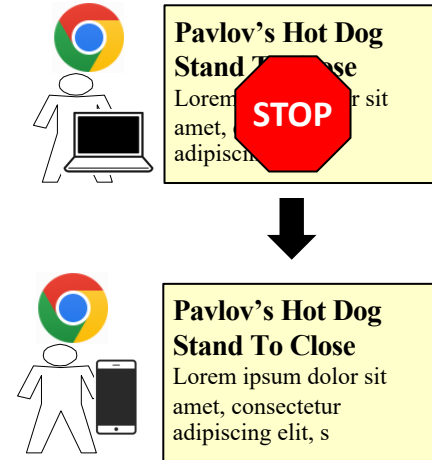
Circumvention: “private” mode or delete cookies



Circumvention: switching browsers



Circumvention: switching devices



Circumvention By Using a Different Cookie

This behavior would appear in the data as shown below.

**Deleting
cookies or
“private” mode**

Date/ time	Cookie ID	IPv4	User ID	URL	Browser	Browser Version	Device	OS	Event
5/12/21 13:44:03	59f3ad22 8a059f2b	50.238. 18.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.418 3.83	PC	Windows	Stop
5/12/21 13:44:31	a88b52b d31064cc	50.238. 18.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.418 3.83	PC	Windows	

**Switching
browsers**

5/12/21 13:44:03	59f3ad228 a059f2b	50.238. 18.210	null	https://www.news.com/lo cal/mayor-elected/	Chrome	85.0.4183. 83	PC	Windows	Stop
5/12/21 13:44:31	a88b52bd 31064cc	50.238. 18.210	null	https://www.news.com/lo cal/mayor-elected/	Edge	85.0.564. 44	PC	Windows	

**Switching
devices**

5/12/21 13:44:03	59f3ad22 8a059f2b	50.238. 18.210	null	https://www.news.com/ local/mayor-elected/	Safari	13.1.2	Mac	macOS	Stop
5/12/21 13:44:31	a88b52b d31064cc	50.238. 18.210	null	https://www.news.com/ local/mayor-elected/	Safari	13.1.2	Phone	iOS	

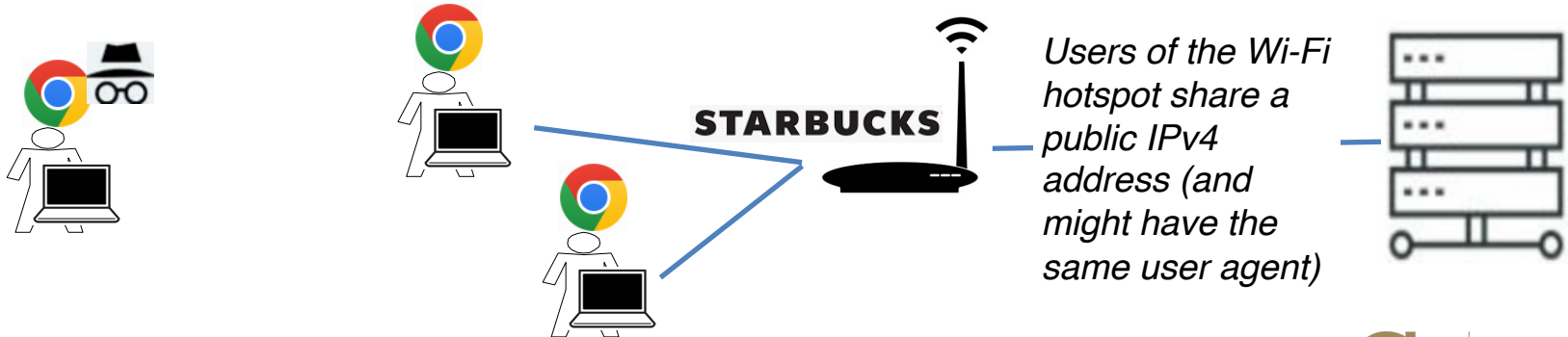
However, how can we tell if this is: 1) a given reader circumventing, or
2) two different readers?

Distinguishing Circumvention from Different Readers

Date/ time	Cookie ID	IPv4	User ID	URL	Browser	Browser Version	Device	OS	Event
5/17/21 13:44:03	59f3ad228 a059f2b	50.238.18.2 10	null	https://www.news.com/local/ mayor-elected/	Chrome	85.0.4183.83	PC	Windows	Stop
5/17/21 13:45:31	a88b52bd3 1064cc	50.238.18.2 10	null	https://www.news.com/local/ mayor-elected/	Chrome	85.0.4183.83	PC	Windows	

Is this a single reader
circumventing by going
into private mode?

Or is it two different readers who share an IP
address (e.g., via coffee shop Wi-Fi) and who both
use Chrome on their laptop?



New Measurement Method

We introduce a novel measurement method for identifying circumvention by using a different cookie.

We rely on a behavioral assumption:

- When a circumventer receives a paywall stop, he still wants to access the article that triggered the stop.



“Triggering” article view



“Possible circumventing” article view

(Some) Measurement Details

We code circumvention when the possible circumventing article view occurs “quickly” after the triggering article view from the same IP address (usually).

a) Via an exponential distribution, model the time between views of the triggering article by readers with a given IP address/browser/device/etc.

Example: 5 different readers, sharing an IP address via a Starbucks Wi-Fi hotspot, accessing www.news.com/local/mayor-elected via Safari on an iPhone.

b) Measure *ElapsedTime* between triggering and possible circumventing article views.

c) Compute the probability that two article views separated by *ElapsedTime* could come from the corresponding exponential distribution. If $p < 0.01 \rightarrow$ circumvention.

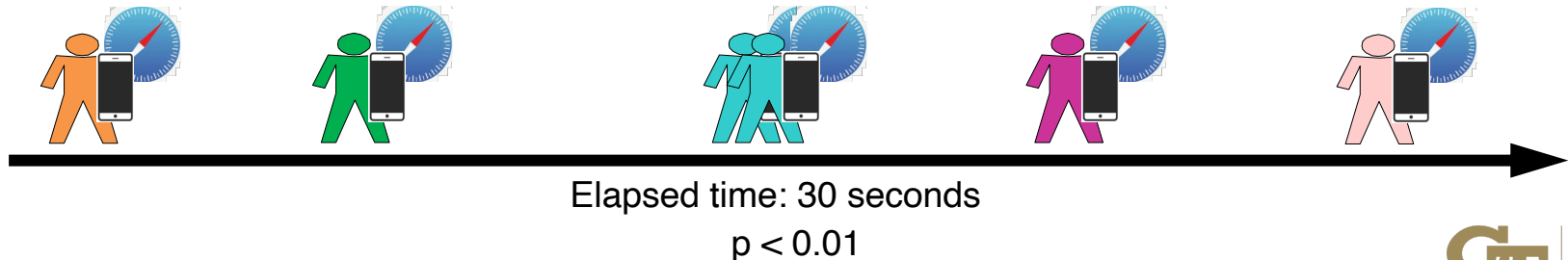


Illustration in the Data

Consider these two cases:

Elapsed
time: 28
seconds

Date/ time	Cookie ID	IPv4	User ID	URL	Browser	Browser Version	Device	OS	Event
5/12/21 13:44:03	59f3ad22 8a059f2b	50.238.1 8.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.418 3.83	PC	Windows	Stop
5/12/21 13:44:31	a88b52b d31064cc	50.238.1 8.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.418 3.83	PC	Windows	

Elapsed
time: 6+
hours

Date/ time	Cookie ID	IPv4	User ID	URL	Browser	Browser Version	Device	OS	Event
5/12/21 13:44:03	59f3ad22 8a059f2b	50.238.1 8.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.418 3.83	PC	Windows	Stop
5/12/21 19:51:31	a88b52b d31064cc	50.238.1 8.210	null	https://www.news.com/ local/mayor-elected/	Chrome	85.0.418 3.83	PC	Windows	

We classify the first as circumvention by deleting cookies / “private” mode and the second as two different readers.

The Key Intuition

The reader's IP address, user agent, etc. are imprecise identifiers.

Thus, we use **the article that the reader circumvents the paywall to access as an additional identifier**, i.e., the missing link.

How Do Readers Respond to the Paywall?

Reader Response	Count	%
Subscribe (or renew)	15,800	0.21%
Login to pre-existing subscriber account	385,970	5.08%
Access non-premium content	1,364,908	17.97%
Probe	950,906	12.52%
Display obstinance	782,180	10.30%
Circumvent	796,897	10.49%
Depart	4,498,110	59.22%

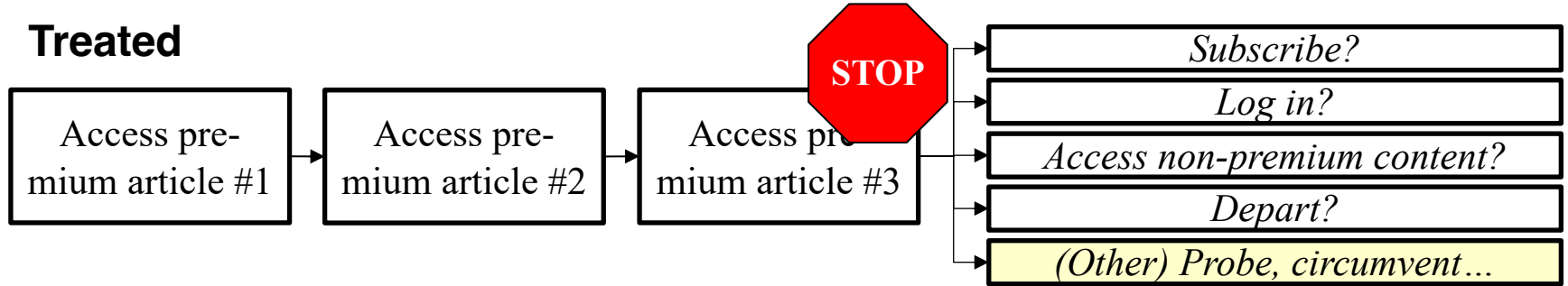
Key takeaway:

- Subscription is rare (0.21%). Circumvention is not (10.49%).

Does the Paywall Make Some Reader Responses More Likely?

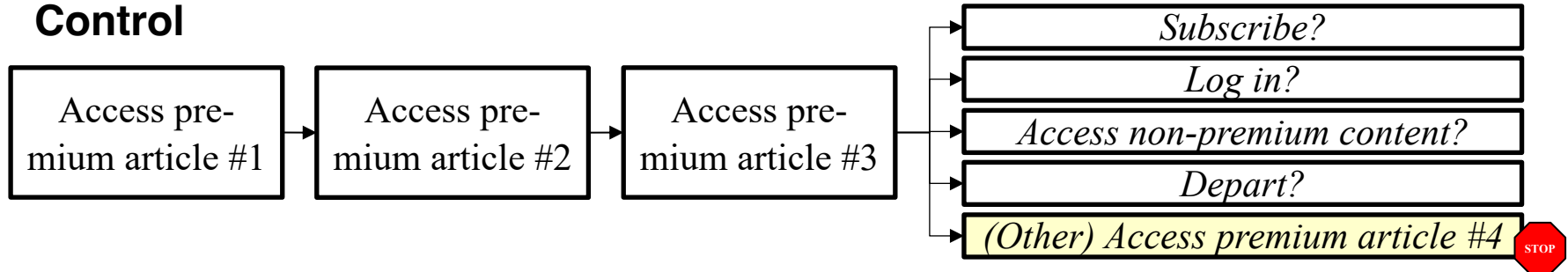
Some readers received the paywall upon accessing their third premium article.

Treated



Some did not. Permits testing the immediate causal effect of the paywall.

Control



Does the Paywall Make Some Reader Responses More Likely?

Variable	Control	Treated
Subscribe	6 (0.00%)	625 (0.10%)
Login	3,598 (0.59%)	58,461 (9.60%)
Access non-premium	141,891 (23.27%)	99,545 (16.34%)
Depart	259,871 (42.62%)	262,649 (43.13%)
Other response	204,428 (33.52%)	187,756 (30.83%)

Key takeaways:

- Almost no one subscribes without hitting the paywall. Presenting the paywall increases the likelihood of subscription by ~100x.
- Presenting the paywall increases the likelihood of login by ~16x.

Factors Predicting Paywall Response: Device

	Subscribe	Circum-vent	Login	Access non-premium	Probe	Obstinance	Depart
PC/Mac	0.0028	0.13	0.06	0.15	0.06	0.04	0.56
Phone	0.0015	0.09	0.04	0.12	0.07	0.07	0.61
Tablet	0.0013	0.09	0.07	0.12	0.07	0.07	0.59

* Predicted probability of each response based on reader device. Results from a discrete choice model.

Key takeaways:

- Readers using a PC/Mac are most likely to respond positively after hitting the paywall:
 - Highest probabilities of subscription, circumvention, and accessing non-premium.

Factors Predicting Paywall Response: Channel

	Subscribe	Circum-vent	Login	Access non-premium	Probe	Obstinance	Depart
Direct	0.0020	0.13	0.06	0.17	0.07	0.05	0.53
Social	0.0013	0.06	0.02	0.02	0.07	0.06	0.77
Search	0.0025	0.07	0.03	0.10	0.06	0.08	0.65

* Predicted probability of each response based on reader channel. Results from a discrete choice model.

Key takeaways:

- Readers accessing the web site directly are most likely to respond positively after hitting the paywall:
 - Highest probabilities of subscription, login, circumvention, and accessing non-premium.
 - Lowest probability of depart.

Factors Predicting Paywall Response: Topic

	Subscribe	Circum-vent	Login	Access non-premium	Probe	Obstinance	Depart
Sports	0.0022	0.14	0.07	0.15	0.05	0.04	0.55
Local	0.0031	0.12	0.06	0.15	0.06	0.04	0.57
Nation	0.0019	0.13	0.07	0.16	0.05	0.03	0.55
Business	0.0026	0.12	0.05	0.14	0.06	0.04	0.58
Opinion	0.0039	0.17	0.08	0.16	0.05	0.04	0.50

* Predicted probability of each response based on article topic. Results from a discrete choice model.

Key takeaways:

- Readers of opinion articles most likely to respond positively after hitting the paywall:
 - Highest probabilities of subscription, login, circumvention, and accessing non-premium.

Key Takeaways

How do readers respond to encountering the paywall?

Few subscribe, most leave, and ~10% circumvent.

How can we measure circumvention, given that it is a surreptitious behavior that is designed **NOT** to be observed?

Use the article that readers circumvent to access as the “missing link”.

What factors predict reader responses to the paywall?

Topic (opinion articles), device (PC/Mac), and channel (direct) most likely to yield positive responses, including circumvention.

Implications

We propose a new method **for identifying when readers circumvent your paywall.**

When you identify circumvention, you can:



Do nothing

Use it as a data point
about that reader;
serve relevant ad

Acknowledge the
reader's savvy and
offer a promotion

Present the
stop page
again

Let us know if we can help!

Thank you!

Corresponding author: eric.overby@scheller.gatech.edu