

The Identity Fragmentation Bias

Tesary Lin* Sanjog Misra[†]

June 29, 2021

[*This is the final working paper version. The published version can be found [here](#).*]

Abstract

Consumers interact with firms across multiple devices, browsers, and machines; these interactions are often recorded with different identifiers for the same consumer. The failure to correctly match different identities leads to a fragmented view of exposures and behaviors. This paper studies the *identity fragmentation bias*, referring to the estimation bias resulted from using fragmented data. Using a formal framework, we decompose the contributing factors of the estimation bias caused by data fragmentation and discuss the direction of bias. Contrary to conventional wisdom, this bias cannot be signed or bounded under standard assumptions. Instead, upward biases and sign reversals can occur even in experimental settings. We then compare several corrective measures, and discuss their respective advantages and caveats.

Keywords: fragmentation, cookies, bias, inference, privacy, measurement

*Boston University Questrom School of Business; tesary@bu.edu;

[†]The University of Chicago Booth School of Business; sanjog.misra@chicagobooth.edu. The authors thank Dean Eckles, Max Farrell, Andrey Fradkin, Yufeng Huang, Garrett Johnson, and Harikesh Nair for helpful comments and suggestions.

1 Introduction

Consumers' digital footprints are increasingly fragmented. A consumer often navigates across multiple websites and devices throughout an online purchase journey. Websites track consumers via cookies, which are text files generated to identify a browser-device combination when consumers first visit a website. However, cookies are browser, device, and site specific. As such, a consumer is often associated with mutually disconnected cookie identifiers across sites and browsers, with each one capturing only a fraction of their ad exposures and behaviors. We call this fracture of identifiers and records **identity fragmentation**.

Websites often deploy several workarounds for the identity fragmentation problem. One solution is using a device ID to identify users. For example, mobile apps frequently obtain IDFA (advertiser ID for Apple devices) and Android's Advertiser ID for tracking. Using a device-level identifier alleviates but does not completely avoid fragmentation, since a user can often navigate through multiple devices within the same journey. For example, Facebook reports that 32% of users who show interest in their mobile ads convert on their desktop.¹ Alternatively, advertising platforms sometimes perform cookie-syncing, which matches several fragmented records across publishers and ad-tech partners. However, cookie syncing can match only up to 60% of fragmented records². Despite the short-comings of device IDs and cookie syncing, they are the preferred workarounds for most websites. Compared to alternatives such as a login wall, these two approaches can often be implemented without users' explicit consent, thus minimizing friction and deterrence of potential customers.

Unfortunately, the advantages of device IDs and synced cookies also make them the target of industry self-regulation that responds to consumers' demand for privacy. Starting with iOS 14, Apple requires that mobile apps seek users' opt-in consent before collecting their device identifiers. On the web end, Chrome follows Firefox and Safari, announcing it will stop supporting 3rd-party cookies starting 2022, also citing privacy concerns. Since all cookie syncing relies on third-party cookies, Chrome's move essentially makes cookie syncing obsolete. As the world becomes more and more attuned to user privacy, identity fragmentation is here to stay.

This paper analytically characterizes how fragmented identifiers bias model estimates and inference, which we call the **identity fragmentation bias**. In its simplest form, the problem manifests as an inability to size the user group. More importantly, causal associations across cookies and devices cannot be reliably measured without a unique identifier. Consider a consumer who sees an ad on her desktop and later buys the advertised product on her phone. If these interactions are taken as independent observations from different users, the causal link between the ad exposure and the purchase is broken. Researchers and firms who use fragmented data may thus estimate the advertising effect with bias.

¹<https://www.facebook.com/business/news/cross-device-measurement>

²<https://www.adexchanger.com/data-driven-thinking/ad-tech-needs-a-shared-id-solution-asap>

We decompose the bias into three components in the context of a linear model. The first component arises because variations in the outcome variable are split. As a result, each fragmented identifier only captures a fraction of the outcome variation, leading to an attenuation bias. The second component emerges from the fracturing of covariates. This fracture prevents the variation of covariates in the mobile record from being associated with outcomes in the laptop record and vice versa, creating an omitted variable bias. The third component is caused by a spurious correlation between fragmented outcomes and covariates. This can occur when a consumer is exposed to more ads on the device than he uses more often (device-level activity bias), or when ads on one device cause a consumer to substitute purchases across devices. The latter two bias components have arbitrary signs and magnitudes. As a result, the direction of the identity fragmentation bias is undetermined without further assumptions. Moreover, this bias does not converge to zero in the limit.

We discuss special cases where the bias can be completely characterized. In particular, we show that when an experiment places equal treatment probabilities (for binary treatments) or intensities (continuous treatments) across sites and devices, the raw estimate is attenuated by a factor equal to the number of fragmented records per user. In this case, we can obtain an unbiased estimate for the true parameter by multiplying the raw estimate with the number of fragments. This finding is closely related to Coey & Bailey (2016), who propose a debiasing estimator for cookie-level estimates under similar conditions.

We show the robustness of this result by examining a number of model perturbations. First, we compare models where the parameters of interest are common across fragments and models where parameters of interest are fragment-specific. Then we examine a setting with an arbitrary number of fragments. We have also analyzed cases where data contain a mixture of fragmented and complete identities. In all cases, the identity fragmentation bias persists and is unsigned without further assumptions.

Guided by the analytic form, we compare several solutions for the identity fragmentation bias. Existing solutions can be categorized into two buckets. The first is identity linking (with cookie syncing being one example), which works by identifying and matching fragmented records that belong to the same user. This approach is widely adopted in the industry. However, even the best identity linking methods only stitch up part of the fragmented records. We show that partially matched data often complicates the bias pattern and prevents estimator bounding and debiasing. The second approach, proposed by Coey & Bailey (2016), adjusts post-experiment estimates based on a set of symmetry assumptions. When the assumptions hold, this approach can effectively debias the average treatment effect estimates without the need to stitch up the raw data.

We then propose a third approach, the **stratified aggregation**. Instead of constructing user-level records, this approach works by constructing user group-level records (aggregation) and refining the group using covariates (stratification). The estimate after stratified aggregation is unbiased, but with a higher variance. By relaxing the need for constructing user-level data,

stratified aggregation can achieve a wider coverage, thereby avoiding the complications caused by identity linking. Compared with the experiment-based adjustment approach, it works without additional assumptions or an experiment. This approach works better when the covariates for constructing the stratified groups are high-dimensional, and when the dataset is large enough to account for the increase in variance.

To illustrate the bias and evaluate stratified aggregation, we conduct an empirical exercise using data from an online retailer. In this setting, consumers interact with the seller across mobile, desktop, and tablet devices; their identities are made known to us by an identity matching vendor. We find that when the data is fragmented at the device level, the ad effect estimates are biased upwards. This upward bias is caused by a dominant device-level activity bias. Using constructed demographic variables, we show that stratified aggregation can remove the bias, but comes at the cost of precision.

Our paper contributes to several strands of literature. The first one examines the role of customer identifiers for delivering accurate insights, including Rossi et al. (1996), Aziz & Telang (2016), Trusov et al. (2016), and Miller & Skiera (2018). While these papers demonstrate the value of unfragmented data for profiling or prediction, we focus on the impact of fragmented data on parameter estimates. More recent papers have documented the economic values of cookie-level data, including Goldfarb & Tucker (2011*b*), Marotta et al. (2019); and Johnson et al. (2020). We complement their analysis by describing how the inability to persistently track users will impact measurement and model estimation.

Second is the literature that recognizes problems from using cookie-level in lieu of individual-level data (Chatterjee et al. 2003, Manchanda et al. 2006, Rutz et al. 2011, Bleier & Eisenbeiss 2015, Hoban & Bucklin 2015, Blake et al. 2016). While these papers mention the possibility of problems, we formally characterize the bias that emerges from identity fragmentation. The analytic form allows us to evaluate and compare different solutions to tackle the fragmentation bias.

Our paper is closely related to Coey & Bailey (2016), who also examine the analytic form of estimation bias caused by identity fragmentation. They show that the use of cookie-level data leads to an attenuation bias in a setting free from confounds such as activity bias (Lewis et al. 2015, Johnson et al. 2017) and cross-channel (device or cookie) substitution (Goldfarb & Tucker 2011*a*). This setting allows them to perform post-experiment adjustment to debias cookie-level estimates. We take a step forward by accounting for activity bias and cross-channel substitution when characterizing the bias. We also relax their assumption that cookie uses are ex-ante symmetric, so that our framework can be generalized to discuss other cross-channel settings.

Several authors have proposed approaches to alleviate the data fragmentation problem. These approaches include data linking (see Abramitzky et al. 2021 and Bailey et al. 2020 for a review), experiment-based adjustments (Coey & Bailey 2016, Koehler et al. 2016), missing data imputation (Novak et al. 2015), and aggregation (Rutz et al. 2011, Blake et al. 2021). Our paper takes another step

forward by comparing existing solutions and proposing a method to improve the last approach. In particular, we propose refining the aggregation using a combination of variables associated with the fragmented identities.³

We organize the rest of the paper as follows. Section 2 illustrates the fragmentation problem using an example. Section 3 formally characterizes the bias, followed by a discussion on bias in experiment settings and generalizations. Section 4 compares several solutions, followed by an empirical application that demonstrate their performances. Section 6 concludes.

2 An Illustrative Example

Consider a setting where each consumer has two devices, a mobile (M) and a desktop (D). Consumers are exposed to ads, and the researcher seeks to estimate the impact of ads on purchases, denoted as β . Suppose the advertising effect does not hinge on which device is used for ad delivery. Table 1(a) presents two consumers who have been exposed to 2 and 4 ads, respectively. Since both consumers purchase equal amounts, advertising has no discernible effect ($\beta = 0$).

Table 1: Estimating Advertising Effects with Fragmented Data: An Example

(a) True Data

Identity	Device	Purchase	Ads	Actual Effect
1	D	1	2	$\beta = 0$
	M			
2	D	1	4	
	M			

(b) Fragmented Data

Observed ID	Device	Purchase	Ads	Est. Effect
1(a)	D	1	2	$\beta = 0.4$
1(b)	M	0	0	
2(a)	D	1	3	
2(b)	M	0	1	

(c) Fragmented Data

Observed ID	Device	Purchase	Ads	Est. Effect
1(a)	D	1	0	$\beta = -0.4$
1(b)	M	0	2	
2(a)	D	1	1	
2(b)	M	0	3	

³This paper is also related to the broad econometrics literature on overcoming unknown network interference, such as Manresa (2016) and Savje et al. (2019).

Now consider the data where identities are fragmented. Since the researcher can no longer associate devices to consumers, she mistakenly assumes that there are four distinct observations. Table 1(b) shows that if consumers see ads on devices that they then purchase on, a spurious positive correlation occurs, creating an upward bias in the estimated advertising effect. In this example, the parameter estimate is $\beta = 0.4$. In other cases, the bias can go in the opposite direction. Suppose instead that consumers see ads more often on their phones but make purchases predominantly on their desktops. Then the estimated advertising effect can be negative, as in panel (c) of Table 1.

This stylized example shows that identity fragmentation can lead to significant biases, which cannot be signed without additional information. In the next section, we formally characterize the identity fragmentation bias and decompose it into three components. We then discuss which feature of fragmented data creates each bias component, and use the finding to guide our discussion on debiasing approaches.

3 Characterizing Identity Fragmentation Bias

We use the term “fragment” to describe any sub-identity that a consumer may have. Fragments can arise across *cookies, devices, channels*, or any other measurement unit across which identities are possibly split. To make the illustration concrete, we focus on the earlier example, where the research goal is estimating the effect of advertising on purchases and where data are fragmented across devices. The analysis and results are not constrained to this example.

To facilitate exposition, we consider estimates under the standard linear model.⁴ In Section 3.1, we assume that each consumer has two devices and that the researcher cannot separate observation units by device types. As such, the researcher does not distinguish ad effect across device types when analyzing the data. We relax this assumption in Section 3.4 and onwards.

3.1 Analysis Framework

The researcher estimates a common effect of advertising exposures on purchases:

$$y = \alpha + x'\beta + \epsilon. \tag{1}$$

Here, y denotes the dollar spent per user; x represents a K -dimensional vector of covariates including advertising exposure, and ϵ is the error component. The parameter of interest is β . Let y_j, x_j denote the corresponding variables on device $j \in \{1, 2\}$. By construction, $y = y_1 + y_2$ and $x = x_1 + x_2$, representing the aggregate spend and exposure levels for a given user. As usual, we also assume $E[\epsilon|x_1, x_2] = 0$.

⁴The sources of bias we present here persists in nonlinear models, though the exact expression of the bias may change.

The un-fragmented (true) data consists of N observations, reflecting N unique consumers whose observations are identically and independently distributed. Let $Y = [y_{(1)}, \dots, y_{(N)}]'$, $X_j = [x'_{(1)j}, \dots, x'_{(N)j}]$, and define $X = [\eta \quad X_1 + X_2]$ where η is a length- N vector of ones. If consumer-level identities were observed, a researcher could obtain an estimate of β by regressing Y on X . When the data is *fragmented*, however, the researcher observes

$$\tilde{Y} \equiv \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \quad \tilde{X} \equiv \begin{bmatrix} \eta & X_1 \\ \eta & X_2 \end{bmatrix}.$$

Without additional information, the researcher treats each device as a different user, assumes that the number of observations is $2N$, and estimates β by regressing \tilde{Y} on \tilde{X} .

An important aspect of identity fragmentation is how purchases occur across devices. We allow consumers to have idiosyncratic preferences over how they use devices for purchases, and capture these preference using a binary variable $s_{(i)}$. In particular, $s_{(i)} = 1$ implies that consumer i completed his purchases on device 1; $s_{(i)} = 0$ means his purchase was made on device 2.⁵ We define \mathbf{s} to be an $N \times N$ diagonal matrix where the i_{th} diagonal is $s_{(i)}$, and stack these matrices to form $S = \begin{bmatrix} \mathbf{s} \\ I - \mathbf{s} \end{bmatrix}$. We can then express the relationship between the fragmented and un-fragmented purchases as

$$\tilde{Y} = SY. \quad (2)$$

We define the expectation of device usage conditional on covariates as $\Lambda_x = \mathbb{E}[\mathbf{s}|(X_1, X_2)]$. Further, we define $W \equiv [I_{N \times N} \quad I_{N \times N}]$, and $\Omega_{(K+1) \times (K+1)} = \text{diag}(1/2, 1_K)$. With this notation, we characterize the relation between the unfragmented covariates (X) and its fragmented version (\tilde{X}) as

$$X = W\tilde{X}\Omega \quad (3)$$

Let $\theta \equiv \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$. The regression estimator using fragmented data can now be written as

$$\hat{\theta} = (\tilde{X}'\tilde{X})^{-1}(\tilde{X}'\tilde{Y}) \quad (4)$$

$$= (\tilde{X}'\tilde{X})^{-1}(\tilde{X}'SY) \quad (5)$$

$$= (\tilde{X}'\tilde{X})^{-1}[\tilde{X}'S(W\tilde{X}'\Omega\theta + \epsilon)] \quad (6)$$

We introduce the following assumptions to exclude other conventional forms of bias, so that the remaining bias is induced *solely* by identity fragmentation:

Assumption 1. $s \perp \epsilon | (x_1, x_2)$;

⁵Making $s_{(i)}$ a continuous variable $\in [0, 1]$ does not change our results. The only reason that we assume $s_{(i)} = 1$ to be binary is that it is unusual for consumers to split their purchases across devices within a shopping journey.

Assumption 2. *Standard OLS assumptions: exogeneity ($\mathbb{E}[\epsilon|X] = 0$), no perfect collinearity ($\mathbb{E}[XX'] > 0$), i.i.d. and homoskedastic errors ϵ .*

Under these assumptions, we can represent the conditional expectation of θ as

$$\mathbb{E}[\hat{\theta}|X_1, X_2] = \Omega\theta + (\tilde{X}'\tilde{X})^{-1} \begin{bmatrix} 0 \\ (X_1 - X_2)'[\Lambda_x - \frac{1}{2}I]\eta\alpha + [-X_1'(I - \Lambda_x)X_1 + X_1'X_2 - X_2'\Lambda_x X_2]\beta \end{bmatrix}. \quad (7)$$

Our main interest is $\hat{\beta}$, the estimate for the slope parameter(s). Let ϑ be the bottom-right sub-matrix of $(\tilde{X}'\tilde{X})^{-1}$, then the conditional mean of estimation bias for β is (see proof in Appendix A.1):

$$\begin{aligned} \mathbb{E}[\hat{\beta}|X_1, X_2] - \beta &= \vartheta \left(\underbrace{(X_1 - X_2)'[\Lambda_x - \frac{1}{2}I]\eta\alpha}_{\Delta_3} + \underbrace{[X_1'X_2]\beta}_{\Delta_2} - \underbrace{[X_1'(I - \Lambda_x)X_1 + X_2'\Lambda_x X_2]\beta}_{\Delta_1} \right) \\ &= \vartheta (\Delta_3 + \Delta_2 + \Delta_1). \end{aligned} \quad (8)$$

3.2 The Bias Components

The multiplier ϑ in Equation (8) is always positive definite. Therefore, the sign of bias in $\hat{\beta}$ is solely determined by the components Δ_1 , Δ_2 , and Δ_3 . Each component represents a specific form of distortion emerging from the split identities: fragmentation of the outcome variable (Δ_1), fragmentation of the exposure variables (Δ_2), and interaction of the fragmented outcome and exposures (Δ_3).

Δ_1 : Purchase Fragmentation

The first bias component results from splits of the outcome variable Y . Each row in the fragmented data only captures a fraction of variation in the original outcome. This type of measurement error manifests itself via the terms $-X_1'(I - \Lambda_x)X_1$ and $-X_2'\Lambda_x X_2$, creating an attenuation bias. To see this mathematically, note that $-I < \vartheta\Delta_1 < 0$. Intuitively, since the number of observations is artificially doubled while the total variation in outcome is constant, there is a direct attenuation on the measured outcome variation, and subsequently, the estimate. In our advertising example, the fact that the average customer spend is lower simply because we think there are $2N$ observations makes advertising seem less effective than it actually is. This component of the fragmentation bias always exists, except in the trivial case when $\beta = 0$.

Δ_2 : Exposure Fragmentation

The second bias component comes from the fragmentation of data on the exposure side. The fragmentation of identities prevents the model from establishing the association between X 's and Y 's *across* fragments. To see this, note that the true covariate is $X = X_1 + X_2$. However, in the fragmented data, some rows have $\tilde{X} = X_1$ and consequently, the variation pertaining to X_2 is omitted; in other rows, the opposite occurs. This omission is represented by the interaction term $X_1'X_2$ in Δ_2 , which is often seen in the characterization of omitted variable bias (see, e.g., Greene 2003 and Wooldridge 2009).

The direction and magnitude of this component depend on $X_1'X_2$. When covariates are normalized to mean-zero before entering the model, $X_1'X_2$ is the correlation of ad exposures across devices. Thus, a stronger, positive (negative) correlation between cross-device ad exposures leads to a positive (negative), larger bias component Δ_2 , the same as what occurs for omitted variable bias. More generally, we see that covariate normalization can change the magnitude of this bias term. In the special case where ad exposures are independent across devices, normalizing covariates to mean zero can eliminate this bias component.

Δ_3 : Spurious Covariance

This term captures a *spurious* correlation between \tilde{Y} and \tilde{X} . Recall that $\tilde{Y} = SY$ and that $E[S|(X_1, X_2)] = \begin{bmatrix} \Lambda_x \\ I - \Lambda_x \end{bmatrix}$. The main element in Δ_3 , the product $(X_1 - X_2)'(\Lambda_x - \frac{1}{2}I)$, is the estimate of $\mathbf{cov}(\tilde{X}, S)$. This is the covariance between advertising exposure and device usage preferences. Two consumer behavioral patterns can cause a nonzero $\mathbf{cov}(\tilde{X}, S)$, thereby creating spurious correlation. The first one is a device-level activity bias. For example, if a consumer is more likely to see ads and buy products on the same device, then $\mathbf{cov}(\tilde{X}, S) > 0$, creating an upward distortion of the estimator. If the ads and purchases are made systematically on different devices, the resultant distortion goes in the opposite direction. The device-level activity bias can impact the estimate even when the device usage preference is independent of exposure levels. That is, even when Λ does not depend on (X_1, X_2) , $\mathbf{cov}(\tilde{X}, S) \neq 0$ as long as $X_1 \neq X_2$ and $\Lambda \neq \frac{1}{2}I$. As long as there are differential device usage preferences and differential exposure levels, this bias term will persist.

Spurious correlation can also occur when consumers substitute their purchases across devices in response to advertising. Suppose a consumer increases the purchases on the phone in response to a mobile ad, but at the same time decrease purchases on his desktop by the same amount. The advertising is ineffective overall. However, the researcher overestimates the advertising effect, because she only observes the positive ad-purchase correlation on the phone and interprets the decreased consumption level on the desktop as the no-ad baseline. From an econometric stand-

point, this bias is caused by the fact that the fragmented measures for purchases, SY and $(1 - S)Y$, have measurement errors correlated with X .

Activity bias and cross-device substitution represent different consumer behaviors and should be tackled differently. Activity bias results from consumers' device usage preferences itself, while cross-device substitution effects arise due to a shift in device usage caused by ad exposures. As we will see later, activity bias can be removed by unstacking the fragmented data (i.e., estimating separate models by device). By contrast, cross-device substitution effects are not eliminated even when data are unstacked.

We note that $\Delta_1 + \Delta_2$ is bounded by $[-\beta, \beta]$. Therefore, in special cases where the baseline outcome α is close to zero, the bias caused by spurious correlation is negligible, and $\widehat{\beta} \in [0, 2\beta]$ (or $[2\beta, 0]$ when $\widehat{\beta}$ is negative). In other words, in this case, the estimate has its sign consistent with the true effect, and also gives us an effect size lower bound $|\beta| > |\widehat{\beta}|/2$. A close-to-zero baseline outcome is more likely to occur for new product sales. On the other hand, in most digital ad effect studies, $\alpha \gg \beta$, and the bias caused by spurious correlation should be taken into account.

3.3 How Does Experiment Help?

In certain conditions, the bias exhibits only attenuation bias with a simple structure. We characterize the condition below:

*The **Symmetric and independent exposures (SIE)** condition is satisfied when both of the following statements hold:*

1. *Symmetry:* $E[x_1] = E[x_2], \text{Var}[x_1^2] = \text{Var}[x_2^2]; x_1 \perp x_2;$
2. *Independence:* $E[\mathbf{s}|X_1, X_2] = \Lambda$ does not depend on (X_1, X_2) .

Under *SIE*, $E[\widehat{\beta}] = \frac{\beta}{2}$ as long as covariates are normalized to mean-zero before entering the model. Put differently: if the researcher can guarantee that the exposure across fragments is symmetric in mean and variance, and that the device preferences do not depend on exposures, then the estimator is attenuated by a factor exactly equal to the number of fragments. In Coey & Bailey (2016), attenuation bias occurs when x_1, x_2 are i.i.d. distributed and a user purchases products using each cookie with equal probability: this is a special case of *SIE*.

We note that both conditions in *SIE* are imperative to guarantee an attenuation bias. In particular, suppose we have an experiment, so that $X_1 \perp X_2$, but different fragment types receive different intensities of the treatment. The direction of fragmentation bias will still be arbitrary in this case. Consider the following numerical example:

$$E[x_1] = 3, E[x_2] = 1, E[x_1^2] = 10, E[x_2^2] = 2; x_1 \perp x_2.$$

Assume that device preferences are independent of exposure so that $\mathbb{E}(s|x) = \lambda_x = \lambda$; and that $\beta > 0, \alpha = 0$. Under this setting, the fragmentation bias is

$$\mathbb{E}[\beta|X_1, X_2] - \beta = \left(2\lambda - \frac{7}{4}\right)\beta.$$

It is straightforward to see that the bias varies with device preferences λ :

$$\begin{aligned}\lambda \in (7/8, 1] &\implies \widehat{\beta} > \beta; \\ \lambda \in (3/8, 7/8) &\implies 0 < \widehat{\beta} < \beta; \\ \lambda \in [0, 3/8) &\implies \widehat{\beta} < 0.\end{aligned}$$

In other words, randomization by itself does not provide additional guarantees regarding the sign of the fragmentation bias.

An experiment that completely randomizes treatments across device types can help ensure that the first condition in *SIE* is satisfied. On the other hand, condition 2 is an assumption on consumer behavior, which is more likely to be satisfied in certain analysis paradigms than others. We return to this point in Section 4 where we discuss experiment-based debiasing solutions.

3.4 Generalizations

Device-Specific Effects

When device types are well-defined (e.g., consumers use phones and computers; in contrast, cookies associated with the same user may all belong to the same type), a researcher may estimate device-specific advertising effects. They may choose one of the two approaches below, which lead to different forms of biases. The first specifies device-specific ad variables while still pooling data across devices. The model is

$$y = \alpha + x'_1\beta_1 + x'_2\beta_2 + \epsilon. \quad (9)$$

The stacked data is now $\tilde{X} \equiv \begin{bmatrix} \eta & X_1 & \emptyset \\ \eta & \emptyset & X_2 \end{bmatrix}$. Let $\widehat{\beta} \equiv [\widehat{\beta}'_1 \quad \widehat{\beta}'_2]'$. Then the bias can be expressed as the following (see proof in Appendix A.2):

$$E \left[\widehat{\beta}|X_1, X_2 \right] - \beta = \vartheta \begin{bmatrix} \underbrace{X'_1(\Lambda_x - \frac{1}{2}I)\eta\alpha}_{\Delta_3} - \underbrace{X'_1(I - \Lambda_x)X_1\beta_1}_{\Delta_1} + \underbrace{X'_2\Lambda_x X_1\beta_2}_{\Delta_2} \\ \underbrace{X'_2(\frac{1}{2}I - \Lambda_x)\eta\alpha}_{\Delta_3} + \underbrace{X'_2(I - \Lambda_x)X_1\beta_1}_{\Delta_2} - \underbrace{X'_2\Lambda_x X_2\beta_2}_{\Delta_1} \end{bmatrix} \quad (10)$$

As before, the estimation bias contains attenuation effect (Δ_1), omitted variable bias (Δ_2), and spurious correlation (Δ_3). However, in contrast to the common-effect model, here consumers' substitution between fragments influences not only the spurious correlation but also the omitted variable bias, as is reflected by the interaction between Λ_x and X_1, X_2 .

The second approach splits data by device type and estimates separate ad effect models. The fragmentation bias in this case is

$$E[\widehat{\beta}|X_1, X_2] - \beta = \begin{bmatrix} (X'_1 X_1)^{-1} \underbrace{(-X'_1(I - \Lambda_x)X_1\beta_1)}_{\Delta_1} + \underbrace{X'_2 \Lambda_x X_1 \beta_2}_{\Delta_2} \\ (X'_2 X_2)^{-1} \underbrace{(-X'_2 \Lambda_x X_2 \beta_2)}_{\Delta_1} + \underbrace{X'_2(I - \Lambda_x)X_1 \beta_1}_{\Delta_2} \end{bmatrix}$$

Activity bias now disappears because the researcher no longer stacks the fragmented data. However, other bias components remain. In particular, bias caused by cross-device substitution may still manifest itself in the omitted-variable-bias term (Δ_2). Moreover, unlike in the common-effect model, here *SIE* no longer guarantees attenuation bias. This is because the bias terms in different devices do not cancel out with each other when data from different devices enter separate models.

Extension to J Fragments

Generalizing our discussion to the case where identities are split into $J \geq 2$ fragments is straightforward. For the common-effect model, the bias is

$$E[\widehat{\beta}|X_1, \dots, X_J] - \beta = \vartheta \left\{ \left[\sum_{j=1}^J X'_j (\Lambda_j - \frac{1}{J}I) \eta \right] \alpha + \left[\sum_{j=1}^J X'_j (\Lambda_j - I) X_j + \sum_{j_1 \neq j_2} X'_{j_1} (\Lambda_{j_1} + \Lambda_{j_2}) X_{j_2} \right] \beta \right\}; \quad (11)$$

where ϑ is the $J \times J$ lower block diagonal of $(\widetilde{X}'\widetilde{X})^{-1}$. Under a more general set of symmetric treatment conditions: $\left\{ E[x_j] = E[x_{j'}], E[x_j^2] = E[x_{j'}^2], x_j \perp x_{j'}, \forall j \neq j' \right\}$, we can show that $E[\beta|x_1, \dots, x_J] = \frac{\beta}{J}$.

We can also extend the results corresponding to the device-specific-effect model to obtain, for the stacked-data version:

$$E[\widehat{\beta}|X_1, \dots, X_J] - \beta = \vartheta \begin{bmatrix} X'_1(\Lambda_1 - \frac{1}{J}I)\eta\alpha + X'_1(\Lambda_1 - I)X_1\beta_1 + \sum_{j \neq 1} X'_1 \Lambda_1 X_j \beta_j \\ \vdots \\ X'_j(\Lambda_j - \frac{1}{J}I)\eta\alpha + X'_j(\Lambda_j - I)X_j\beta_j + \sum_{j \neq j} X'_j \Lambda_j X_j \beta_j \end{bmatrix}; \quad (12)$$

and for the split-sample version:

$$E[\widehat{\beta}|X_1, \dots, X_J] - \beta = \begin{bmatrix} (X_1'X_1)^{-1}(X_1'(\Lambda_1 - I)X_1\beta_1 + \sum_{j \neq 1} X_1'\Lambda_1 X_j\beta_j) \\ \vdots \\ (X_J'X_J)^{-1}(X_J'(\Lambda_J - I)X_J\beta_J + \sum_{j \neq J} X_J'\Lambda_J X_j\beta_j) \end{bmatrix}. \quad (13)$$

Results associated with the bias are analogous to the two-device version.

3.5 Other Extensions

While associated results are not reported, we have considered settings where the researcher is interested in other model forms, including treatment-effect type estimators and nonlinear models (e.g., logit). We have also considered cases where data may exhibit false matching, partial matching, or probabilistic matching.⁶ Our general finding across these models, settings, and specifications is that (a) the identity fragmentation bias continues to exist; (b) the bias structure is driven by some combination of the three components aforementioned; (c) in most cases the sign of the bias cannot be easily determined.

4 Solution Comparison

In this section, we compare three solutions to identity fragmentation bias: identity linking, experiment-based estimator adjustment, and stratified aggregation. Below we discuss their respective strengths and limitations.

4.1 Identity Linking

Identity linking is often the go-to solution for practitioners. This approach aims to construct individual-level data from the fragmented ones, either by matching fragments that share the same characteristics (probabilistic matching) or by exchanging credentials associated with the same user (deterministic matching).

It is not hard to see the intuitive appeal of this approach. However, as the advertising industry responds to privacy concerns, many previously prevalent linking methods will become obsolete. For example, many probabilistic matching methods rely on third-party cookies, which will no longer work once Chrome stops their third-party cookie support.⁷ Other methods that rely on first-party data requires consent from consumers for data to be shared. As a result, these methods can only link part of the fragmented records in practice.

⁶Many of these analysis may be added to the paper in subsequent versions. They are available from the authors upon request.

⁷<https://headerbidding.co/universal-id-adtech/>

In some cases, partially linked data can exacerbate the fragmentation bias. Suppose the pooled data contain a proportion r of fragmented users. The resulting estimator is the weighted average of the “pure” estimators (Durbin 1953, Theil & Goldberger 1961):

$$\widehat{\theta}^m = \omega \widehat{\theta}^f + (I - \omega) \widehat{\theta}^l, \quad \text{where} \quad \omega = (r \widetilde{X}' \widetilde{X} + (1 - r) X' X)^{-1} r \widetilde{X}' \widetilde{X},$$

where $\widehat{\theta}^f$ is the estimator using only the fragmented data, while $\widehat{\theta}^l$ is the estimator obtained from the unfragmented data alone. It follows that $E[\widehat{\theta}^m | X] - \theta = \omega(E[\widehat{\theta}^f | X] - \theta)$. The problems with this estimator are that since ω is a matrix, (i) ω does not necessarily increase “monotonically” with r ; (ii) ω redistributes the bias in each element inside the vector $E[\widehat{\theta}^f | X] - \theta$. In simulations, we see that with intermediate ranges of r , $\widehat{\theta}^m$ is often further away from θ , and sometimes the bias takes the reversed sign compared to $\widehat{\theta}^f$. The potential sign reversal and the amplified bias magnitude makes it even harder to bound the bias term.

4.2 Experiment-Based Estimator Adjustment

In Section 3.3 and 3.4, we show that when *SIE* holds, the slope estimator in a common-effect model is attenuated by the number of fragments per user. Thus, we can construct an unbiased estimator, $J\widehat{\beta}$, by multiplying the original estimator with the number of fragments. The confidence interval of this adjusted estimator is at least \sqrt{J} times the confidence interval of the estimator with unfragmented data. This lack of power is caused by the loss of variation created by fragmentation.

To use this debiasing approach, a key question is when *SIE* holds. The symmetry condition in *SIE* can be satisfied by running an experiment that completely randomizes treatments (e.g., ad exposures) across devices. However, due to activity bias, the independence condition is likely to be violated when the researcher estimates the actual treatment effect. Thus, we recommend researchers focus on intent-to-treat: in this case, both conditions in *SIE* hold as long as a completely randomized experiment takes place. We also note that when users each has a different number of fragments, the estimator will assign more weights to users with more fragments. In practice, these are likely users with more activities.

4.3 Stratified Aggregation

A third solution is to aggregate fragmented units into higher-level groups, such that all fragmented identities of a consumer are collapsed within the aggregation. A simple form of aggregation is analyzing data at the geographic level, such as county, city, or stores (Joo et al. 2013, Kalyanam et al. 2018, Blake et al. 2021). While simple aggregation provides robustness, it significantly reduces statistical power and restricts the ability to examine heterogeneity.

To alleviate the power problem, we propose stratified-aggregation. This approach uses observable user characteristic combinations to construct more granular bins, such that all fragments of a given user still fall into the same bin. For example, we can define bins by collapsing fragments that have the same zipcode \times gender \times age \times education combination. The idea of using combinations of non-unique user characteristics to construct unique values for matching is similar to the technique used in probabilistic matching. However, unlike probabilistic matching, stratified aggregation does not require that matching be conducted at the individual level. In doing so, it allows for a more complete match. Given the caveats of partially matched data, we believe that stratified aggregation is more desirable.

The granularity (and in turn, the effective sample size and power) increases with the number of observable user characteristics. For example, Rocher et al. (2019) show that for the US population, a group constructed from 3 demographic variables consists of a unique individual 58% of the time; with a 4th variable, the probability of the group consisting of one individual increases to 80%. Thus, stratified aggregation is more effective when there are many observable user characteristics (columns) and when the number of observations is large (rows). We caution the readers that stratified aggregation may not perform as well when the covariates used for matching are measured with error (e.g., Neumann et al. 2019).

As with experiment-based adjustment, stratified aggregation gives different weights to different users. In this case, users in smaller bins are weighted more. These users are likely consumers with more frequent activities or with rarer attributes. One way to correct for the uneven weights is by obtaining the distribution statistics of the variables used for constructing the strata, so that one can assign weights proportional to the strata sizes.

Summary. Although identity linking is popular among practitioners, results based on partially linked data can be misleading. We recommend that research that uses identity linking services obtain metrics for the match rate, and interpret their estimates with caution. Experiment-based adjustment requires an experiment that completely randomizes treatment across fragment types (or exogenous variation that achieves the same effect), and only delivers a valid intent-to-treat effect under a common-effect model. Nevertheless, it is simple to implement when all the conditions hold. Stratified aggregation requires the least assumptions, and can be applied when the model of interest is nonlinear or even structural. However, it comes at the cost of statistical power.

5 Empirical Application

The goal of this application is to show the magnitude of identity fragmentation bias in a realistic setting and to demonstrate the performance of stratified aggregation. We first describe the institu-

tional background, and present data patterns that indicate the existence of identity fragmentation bias, then proceed to the estimation result discussion.

5.1 The Context

We focus on the problem of digital *attribution*: assigning credit for an outcome of interest (e.g., visits, engagements, conversions) to a variety of marketing touch-points. Satisfactory solution to the attribution problem relies on the ability to estimate the effect of marketing interventions on the outcomes. These estimators are often calibrated on cookie-level or device-level data. Therefore, problems created by identity fragmentation are relevant in this context.

Our data comes from an online retailer of durable goods. Our goal is to estimate the impact of online advertising on customer engagement, measured by visits to the firm’s website. The data has 391,195 observations and includes the following variables: engagement (Y); display ads, search ads, and social media ads (X).

The original data contains individual-level observations, as well as the devices where engagements and advertising exposures take place. These devices are classified as desktop, mobile phone, and tablets. We construct the fragmented data by dividing each observation into three fragments, each corresponding to a specific device type. This data is observational and we warn the reader from interpreting effects as causal (e.g., targeting bias can skew the estimates upward compared to the true causal effects).

5.2 Data Descriptives

Three types of bias components contribute to the identity fragmentation bias: attenuation bias, omitted variable bias, and spurious correlation. The attenuation bias induced by measurement error on Y always exists in fragmented data. Omitted variable biases will exist if X_j ’s are correlated with each other. In our setting, advertising exposures on different devices are not significantly correlated with each other, so omitted variable bias is likely to be secondary. Therefore, we focus our attention on examining the evidence of activity bias and device substitution effects, the two sources of spurious correlation.

Activity bias exists if the average tendency of engagement on each device is positively correlated with advertising exposure on each device. Figure 1 shows the distribution of ads on different devices: it is heavy on mobile phones, moderate on desktops, and minimal for tablets. The distribution of engagements (Figure 2) shows a similar pattern across devices. The resemblance of the two histograms suggests the existence of activity bias.

Device substitution during purchases in response to advertising exposure will also contribute to the bias. If there is no device substitution ($s_j \perp (X_1, \dots, X_J)$), then $Corr(Y_j, X_j) =$

Figure 1: Ad Exposure across Devices

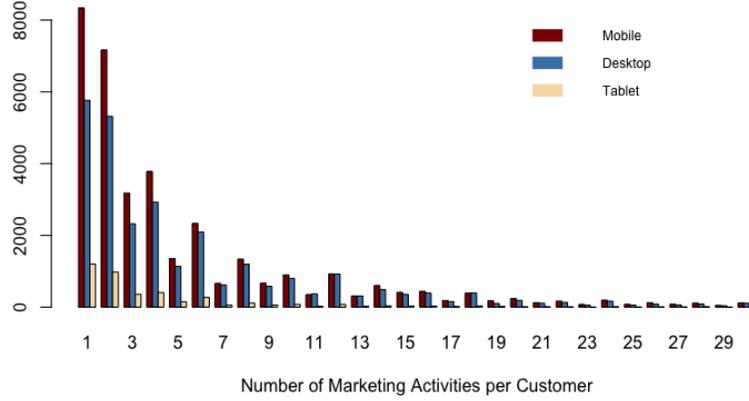
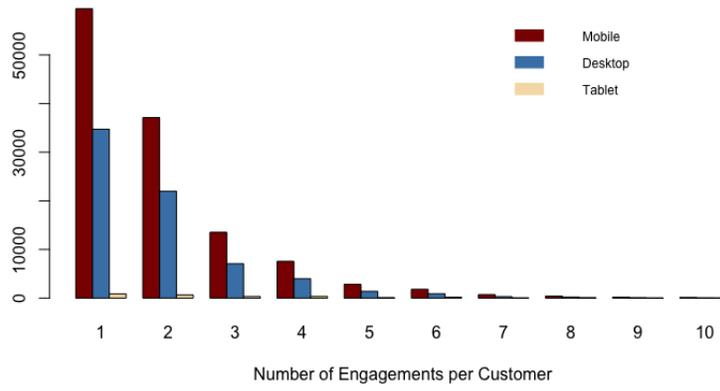


Figure 2: Engagement Distribution across Devices



$\Lambda_j \cdot Corr(Y, X_{j'}), \forall j, j'$. When $J = 3$, this means that

$$\begin{aligned} & Corr(Y_1, X_1) : Corr(Y_2, X_1) : Corr(Y_3, X_1) = Corr(Y_1, X_2) : Corr(Y_2, X_2) : Corr(Y_3, X_2) \\ & = Corr(Y_1, X_3) : Corr(Y_2, X_3) : Corr(Y_3, X_3) = \Lambda_1 : \Lambda_2 : \Lambda_3. \end{aligned}$$

In other words, if device substitution is absent, we should see different columns of the correlation matrix proportional to each other (same for rows). Table 2 shows that this is not the case.

5.3 Estimation Results

Table 3 compares the estimates using true data with those using fragmented data. Across all specifications, identity fragmentation leads to inflated estimates. In the common-effect specification, the magnitude of inflation ranges from 40% (search ad effect) to 88% (display ad effect). Identity fragmentation also leads to narrower confidence intervals compared to the true estimates: this is caused by the fact that the number of “observations” is tripled in the fragmented data. The combination of inflated point estimates and shrunk confidence interval leads to zero overlap

Table 2: Correlations Between Purchase Patterns and Advertising Exposure across Devices

(a) Search			
	Search _m	Search _d	Search _t
Y _m	0.2057	-0.0439	-0.0128
Y _d	-0.0446	0.1914	-0.0172
Y _t	-0.0201	-0.0117	0.2422
(b) Social			
	Social _m	Social _d	Social _t
Y _m	0.0642	-0.0229	-0.0046
Y _d	-0.0163	0.0949	-0.0039
Y _t	-0.0064	-0.0079	0.0715
(c) Display			
	Display _m	Display _d	Display _t
Y _m	0.1363	-0.0756	-0.0192
Y _d	-0.0488	0.2147	-0.0131
Y _t	-0.0224	-0.0218	0.1518

between the fragmented and true estimates. In the device-specific-effect model, the inflation in estimates ranges from 28% (Social on Tablet) to 91% (Display on Mobile). These upward biases could severely mislead marketing investment decisions if ignored.

5.4 Stratified Aggregation

We now examine a proposed solution for the fragmentation bias. Since the original dataset does not contain customer demographics or other variables to aggregate around, we simulate the demographic variables and use them to augment the original data. For each true individual, we assign an *metropolitan statistical area* uniformly drawn from one of the 48 districts; *age* drawn uniformly within the [18, 82] range; and *income* drawn from one of 10 discrete buckets.

Figure 3 shows the results of the stratified aggregation exercise. Although the 95% confidence intervals for the aggregated estimates are larger in most cases, they contain the confidence intervals from the unfragmented data. In the device-specific effect model, the stratified aggregation estimates for the tablet ad effects are very noisy, since both purchases and ad exposures on the tablet are scarce. Even so, the results are superior to those obtained from fragmented data, which often do not contain the estimates using individual-level data.

Table 3: Estimates Using Individual Level vs. Fragmented Data

(a) Common Effects Model

	True Est	True SE	Fragmented Est	Fragmented SE	Bias Ratio ($[3]/[1]$)
Intercept	0.9540	0.0024	0.3008	0.0009	0.3153
Search	0.2584	0.0025	0.3626	0.0016	1.4032
Social	0.1439	0.0039	0.2157	0.0026	1.4991
Display	0.0428	0.0007	0.0806	0.0005	1.8816

(b) Device-Specific Effects Model

	True Est	True SE	Fragmented Est	Fragmented SE	Bias Ratio ($[3]/[1]$)
Intercept	0.9535	0.0024	0.3006	0.0009	3.1720
Search.m	0.2564	0.0031	0.3699	0.0020	1.4430
Social.m	0.1191	0.0047	0.1859	0.0031	1.5603
Display.m	0.0396	0.0010	0.0759	0.0006	1.9160
Search.d	0.2741	0.0052	0.3733	0.0034	1.3622
Social.d	0.1985	0.0075	0.2836	0.0049	1.4285
Display.d	0.0456	0.0010	0.0849	0.0007	1.8597
Search.t	0.2394	0.0076	0.2979	0.0050	1.2445
Social.t	0.2516	0.0280	0.3238	0.0183	1.2868
Display.t	0.0449	0.0038	0.0813	0.0025	1.8116

6 Discussion and Future Directions

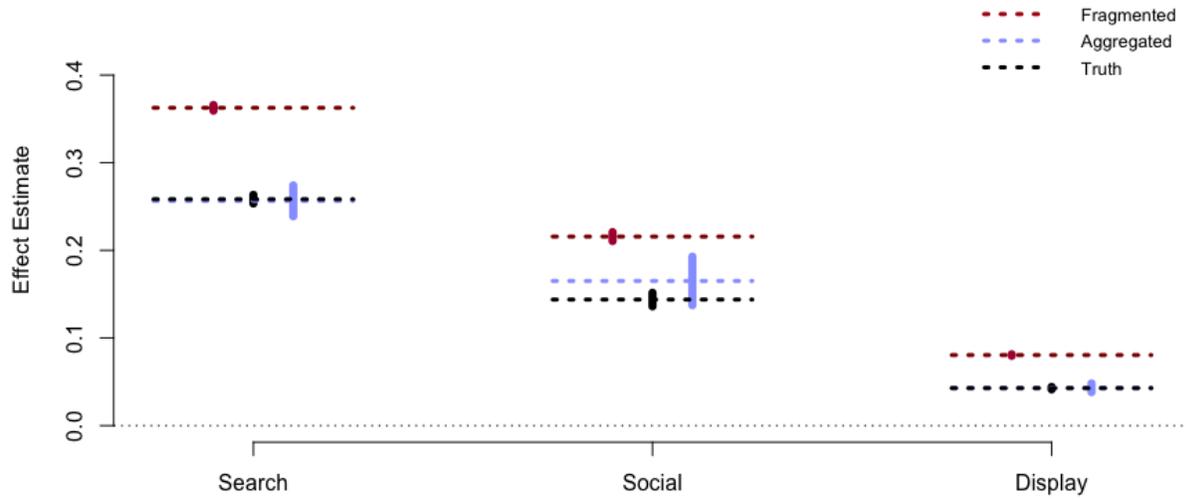
Using data with fragmented user identifiers leads to a persistent bias in estimators. We analytically characterize the bias and show that this bias can have an arbitrary sign. Based on this analytic characterization, we compare three debiasing solutions. Despite its popularity, identity linking can exacerbate the bias and complicate the effort of bounding the estimate when linking is incomplete. Experiment-based estimator adjustment relies on a set of strong assumptions and practicality requirements, but is simple and effective when these requirements are met. Stratified aggregation requires the least assumptions, and can be applied to obtain unbiased estimates for a more generalized set of models. However, it comes at a cost of statistical power.

Although we base our formal analysis on a linear-model setup, the results can also apply to generalized linear models, including discrete choice models and Poisson regressions. For example, one can rewrite a logistic regression as a linear regression of log odds on covariates. The attenuation bias is more likely to occur in discrete choice models, since estimated magnitudes are relative to the residual term. When the model is incorrectly specified, the residual term tends to be larger, and we expect the estimators to exhibit stronger attenuation.

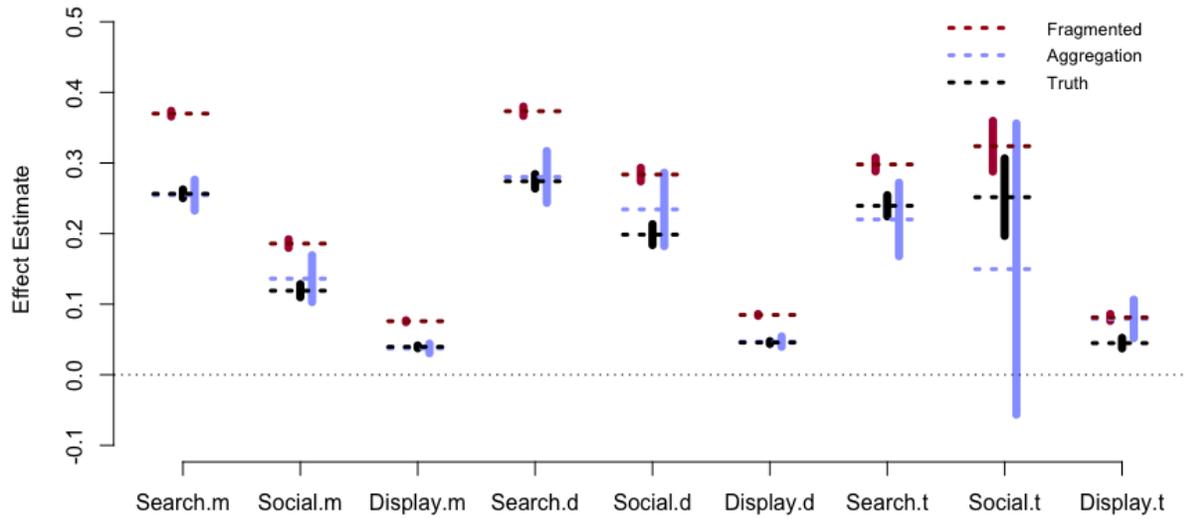
We look forward to future research that further investigates solutions to identity fragmentation bias. Without these solutions, the post-cookie world can render data analytics difficult or even impossible for most firms, while a few established firms who can obtain complete user data using

Figure 3: Estimator Comparison across Different Data

(a) Common Effects Model



(b) Device-Specific Effects Model



Note: The horizontal dashed lines indicate the estimated effect sizes; the vertical lines are the 95% confidence region.

a login wall can have a stronger incumbent advantage. By characterizing the bias and discussing existing solutions, we hope to contribute to this discussion and provide a more even playing field in the cookieless landscape.

References

- Abramitzky, R., Boustan, L. P., Eriksson, K., Feigenbaum, J. J. & Pérez, S. (2021), 'Automated linking of historical data', *Journal of Economic Literature* .
- Aziz, A. & Telang, R. (2016), 'What is a cookie worth?', *Available at SSRN 2757325* .
- Bailey, M., Cole, C., Henderson, M. & Massey, C. (2020), 'How well do automated methods perform in historical samples? evidence from new ground truth', *Journal of Economic Literature* **58**(4), 997–1044.
- Blake, T., Moshary, S., Sweeney, K. & Tadelis, S. (2021), 'Price salience and product choice', *Marketing Science* .
- Blake, T., Nosko, C. & Tadelis, S. (2016), Returns to consumer search: Evidence from ebay, in 'Proceedings of the 2016 ACM Conference on Economics and Computation', ACM, pp. 531–545.
- Bleier, A. & Eisenbeiss, M. (2015), 'Personalized online advertising effectiveness: The interplay of what, when, and where', *Marketing Science* **34**(5), 669–688.
- Chatterjee, P., Hoffman, D. L. & Novak, T. P. (2003), 'Modeling the clickstream: Implications for web-based advertising efforts', *Marketing Science* **22**(4), 520–541.
- Coe, D. & Bailey, M. (2016), People and cookies: Imperfect treatment assignment in online experiments, in 'Proceedings of the 25th International Conference on World Wide Web', International World Wide Web Conferences Steering Committee, pp. 1103–1111.
- Durbin, J. (1953), 'A note on regression when there is extraneous information about one of the coefficients', *Journal of the American Statistical Association* **48**(264), 799–808.
- Goldfarb, A. & Tucker, C. (2011a), 'Search engine advertising: Channel substitution when pricing ads to context', *Management Science* **57**(3), 458–470.
- Goldfarb, A. & Tucker, C. E. (2011b), 'Privacy regulation and online advertising', *Management science* **57**(1), 57–71.
- Greene, W. H. (2003), *Econometric analysis*, Pearson Education India.
- Hoban, P. R. & Bucklin, R. E. (2015), 'Effects of internet display advertising in the purchase funnel: Model-based insights from a randomized field experiment', *Journal of Marketing Research* **52**(3), 375–393.
- Johnson, G. A., Lewis, R. A. & Nubbemeyer, E. I. (2017), 'Ghost ads: Improving the economics of measuring online ad effectiveness', *Journal of Marketing Research* **54**(6), 867–884.
- Johnson, G. A., Shriver, S. K. & Du, S. (2020), 'Consumer privacy choice in online advertising: Who opts out and at what cost to industry?', *Marketing Science* **39**(1), 33–51.
- Joo, M., Wilbur, K. C., Cowgill, B. & Zhu, Y. (2013), 'Television advertising and online search', *Management Science* **60**(1), 56–73.
- Kalyanam, K., McAteer, J., Marek, J., Hodges, J. & Lin, L. (2018), 'Cross channel effects of search engine advertising on brick & mortar retail sales: Meta analysis of large scale field experiments on google. com', *Quantitative Marketing and Economics* **16**(1), 1–42.
- Koehler, J., Skvortsov, E., Ma, S. & Liu, S. (2016), 'Measuring cross-device online audiences'.
URL: <http://pub-tools-public-publication-data.storage.googleapis.com/pdf/45353.pdf>
- Lewis, R., Rao, J. M. & Reiley, D. H. (2015), Measuring the effects of advertising: The digital frontier, in 'Economic Analysis of the Digital Economy', University of Chicago Press, pp. 191–218.

- Manchanda, P., Dubé, J.-P., Goh, K. Y. & Chintagunta, P. K. (2006), 'The effect of banner advertising on internet purchasing', *Journal of Marketing Research* **43**(1), 98–108.
- Marotta, V., Abhishek, V. & Acquisti, A. (2019), Online tracking and publishers' revenues: An empirical analysis, in 'Workshop on the Economics of Information Security'.
- Miller, K. & Skiera, B. (2018), 'Economic loss of cookie lifetime restrictions', *Available at SSRN 3046254* .
- Neumann, N., Tucker, C. E. & Whitfield, T. (2019), 'Frontiers: How effective is third-party consumer profiling? evidence from field studies', *Marketing Science* **38**(6), 918–926.
- Novak, J., Feit, E. M., Jensen, S. & Bradlow, E. (2015), 'Bayesian imputation for anonymous visits in crm data', *Available at SSRN 2700347* .
- Rocher, L., Hendrickx, J. M. & De Montjoye, Y.-A. (2019), 'Estimating the success of re-identifications in incomplete datasets using generative models', *Nature communications* **10**(1), 1–9.
- Rossi, P. E., McCulloch, R. E. & Allenby, G. M. (1996), 'The value of purchase history data in target marketing', *Marketing Science* **15**(4), 321–340.
- Rutz, O. J., Trusov, M. & Bucklin, R. E. (2011), 'Modeling indirect effects of paid search advertising: Which keywords lead to more future visits?', *Marketing Science* **30**(4), 646–665.
- Theil, H. & Goldberger, A. S. (1961), 'On pure and mixed statistical estimation in economics', *International Economic Review* **2**(1), 65–78.
- Trusov, M., Ma, L. & Jamal, Z. (2016), 'Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting', *Marketing Science* **35**(3), 405–426.
- Wooldridge, J. M. (2009), 'Omitted variable bias: the simple case', *Introductory Econometrics: A Modern Approach*. Mason, OH: Cengage Learning pp. 89–93.

Appendix A Proof for Identity Fragmentation Bias

A.1 Common-Effect Model

Under the assumptions laid out in the paper, we have $\epsilon \perp (s, x_1, x_2) \Rightarrow E[g(x_1, x_2, s)\epsilon] = 0, \forall g$. Therefore:

$$\begin{aligned} E[\widehat{\theta}|X_1, X_2] &= (\widetilde{X}'\widetilde{X})^{-1} \cdot \left(\widetilde{X}' \cdot \begin{bmatrix} \Lambda_x \\ I - \Lambda_x \end{bmatrix} W \cdot \widetilde{X} \right) \cdot \Omega\theta \\ &= \Omega\theta + (\widetilde{X}'\widetilde{X})^{-1} \cdot \left(\widetilde{X}' \cdot \begin{bmatrix} (\Lambda_x - I)I & \Lambda_x \\ (I - \Lambda_x)I & -\Lambda_x \end{bmatrix} \cdot \widetilde{X} \right) \cdot \Omega\theta. \end{aligned} \quad (\text{A.1})$$

Since $\widetilde{X}' \cdot \begin{bmatrix} (\Lambda_x - I)I & \Lambda_x \\ (I - \Lambda_x)I & -\Lambda_x \end{bmatrix} = \begin{bmatrix} 0 \\ (X_1 - X_2) \end{bmatrix}' \cdot \begin{bmatrix} \Lambda_x - I & \Lambda_x \end{bmatrix}$, equation (A.1) can be written as

$$\begin{aligned} E[\widehat{\theta}|X_1, X_2] &= \Omega\theta + (\widetilde{X}'\widetilde{X})^{-1} \cdot \left(\begin{bmatrix} 0 \\ (X_1 - X_2) \end{bmatrix}' \cdot \begin{bmatrix} \Lambda_x - I & \Lambda_x \end{bmatrix} \cdot \begin{bmatrix} \eta & X_1 \\ \eta & X_2 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2}\alpha \\ \beta \end{bmatrix} \right) \\ &= \Omega\theta + (\widetilde{X}'\widetilde{X})^{-1} \cdot \left(\begin{bmatrix} 0 & 0 \\ (X_1 - X_2)'(2\Lambda_x - I)\eta & (X_1 - X_2)'(\Lambda_x - I)X_1 + \Lambda_x X_2 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2}\alpha \\ \beta \end{bmatrix} \right) \\ &= \Omega\theta + (\widetilde{X}'\widetilde{X})^{-1} \begin{bmatrix} 0 \\ (X_1 - X_2)'(\Lambda_x - \frac{1}{2}I)\eta\alpha + (X_1'((\Lambda_x - I)X_1 + X_1'X_2 - X_2'\Lambda_x X_2)\beta) \end{bmatrix}. \end{aligned} \quad (\text{A.2})$$

To get $E[\beta|X_1, X_2] - \beta$, only the bottom-right block entry of $(\widetilde{X}'\widetilde{X})^{-1}$ needs to be calculated. Let $\widetilde{X}'\widetilde{X} = \begin{bmatrix} 2N & \eta'(X_1 + X_2) \\ (X_1' + X_2')\eta & X_1'X_1 + X_2'X_2 \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ and $(\widetilde{X}'\widetilde{X})^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & \vartheta \end{bmatrix}$, then

$$\begin{aligned} E[\beta|X_1, X_2] &= \beta + B_{21} \cdot 0 + \vartheta \cdot \left((X_1 - X_2)'(\Lambda_x - \frac{1}{2}I)\eta\alpha + (X_1'(\Lambda_x - I)X_1 + X_1'X_2 - X_2'\Lambda_x X_2)\beta \right) \\ \Rightarrow E[\beta|X_1, X_2] - \beta &= \vartheta \cdot \left((X_1 - X_2)'(\Lambda_x - \frac{1}{2}I)\eta\alpha + (X_1'(\Lambda_x - I)X_1 + X_1'X_2 - X_2'\Lambda_x X_2)\beta \right), \end{aligned} \quad (\text{A.3})$$

where

$$\vartheta = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} = \left[X_1'X_1 + X_2'X_2 - \frac{(X_1' + X_2')\eta \cdot \eta'(X_1 + X_2)}{2N} \right]^{-1}. \quad (\text{A.4})$$

Note that ϑ^{-1} is always positive definite. Moreover,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \vartheta^{-1} = \text{Var}[X_1] + \text{Var}[X_2] + \frac{1}{2}(E[X_1] - E[X_2])^2$$

is also positive definite.

A.2 Device-Specific-Effect Model

With a device-specific-effect model, we use the same notation to represent slightly different objects.

The first change involves the covariates: $\tilde{X} \equiv \begin{bmatrix} \eta & X_1 & \emptyset \\ \eta & \emptyset & X_2 \end{bmatrix}$. Second, the dimension of Ω changes from $k + 1$ to $2k + 1$; as before, the first element of Ω is $1/2$ and the rest are 1. With the same notation, Equation (A.1) in Section A.1 still holds. Now

$$\begin{aligned} \tilde{X}' \cdot \begin{bmatrix} (\Lambda_x - I)I & \Lambda_x \\ (I - \Lambda_x)I & -\Lambda_x \end{bmatrix} \cdot \tilde{X} &= \tilde{X}' \cdot \begin{bmatrix} I \\ -I \end{bmatrix} \cdot \begin{bmatrix} \Lambda_x - I & \Lambda_x \end{bmatrix} \cdot \tilde{X} \\ &= \begin{bmatrix} \emptyset & \emptyset & \emptyset \\ X_1'(2\Lambda_x - I)\eta & X_1'(\Lambda_x - I)X_1 & X_1'\Lambda_x X_2 \\ X_2'(I - 2\Lambda_x)\eta & X_2'(I - \Lambda_x)X_1 & -X_2'\Lambda_x X_2 \end{bmatrix}. \end{aligned} \quad (\text{A.5})$$

Combining Equation (A.5) with Equation (A.1),

$$\begin{aligned} E[\hat{\theta}|X_1, X_2] &= \Omega\theta + (\tilde{X}'\tilde{X})^{-1} \cdot \begin{bmatrix} \emptyset & \emptyset & \emptyset \\ X_1'(2\Lambda_x - I)\eta & X_1'(\Lambda_x - I)X_1 & X_1'\Lambda_x X_2 \\ X_2'(I - 2\Lambda_x)\eta & X_2'(I - \Lambda_x)X_1 & -X_2'\Lambda_x X_2 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2}\alpha \\ \beta_1 \\ \beta_2 \end{bmatrix}, \\ &= \Omega\theta + (\tilde{X}'\tilde{X})^{-1} \cdot \begin{bmatrix} 0 \\ X_1'(\Lambda_x - \frac{1}{2}I)\eta\alpha & X_1'(\Lambda_x - I)X_1\beta_1 & X_1'\Lambda_x X_2\beta_2 \\ X_2'(\frac{1}{2}I - \Lambda_x)\eta\alpha & X_2'(I - \Lambda_x)X_1\beta_1 & -X_2'\Lambda_x X_2\beta_2 \end{bmatrix}, \end{aligned}$$

thus

$$E[\hat{\beta}] - \beta = \vartheta \cdot \begin{bmatrix} X_1'(\Lambda_x - \frac{1}{2}I)\eta\alpha & X_1'(\Lambda_x - I)X_1\beta_1 & X_1'\Lambda_x X_2\beta_2 \\ X_2'(\frac{1}{2}I - \Lambda_x)\eta\alpha & X_2'(I - \Lambda_x)X_1\beta_1 & -X_2'\Lambda_x X_2\beta_2 \end{bmatrix}.$$

Here as before, ϑ is the bottom-right diagonal matrix of $(\tilde{X}'\tilde{X})^{-1}$; now it has dimension $2k \times 2k$.

Note that $\tilde{X}'\tilde{X} = \begin{bmatrix} 2N & \eta'X_1 & \eta'X_1 \\ X_1'\eta & X_1'X_1 & \emptyset \\ X_2'\eta & \emptyset & X_2'X_2 \end{bmatrix}$. Let $A_{11} \equiv 2N$, $A_{12} = \begin{bmatrix} \eta'X_1 & \eta'X_1 \end{bmatrix}$, $A_{21} = \begin{bmatrix} X_1'\eta \\ X_2'\eta \end{bmatrix}$, and

$A_{22} = \begin{bmatrix} X_1'X_1 & \emptyset \\ \emptyset & X_2'X_2 \end{bmatrix}$. Then:

$$\vartheta = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} = \begin{bmatrix} X_1'(I - \frac{\eta\eta'}{2N})X_1 & -X_1' \cdot \frac{\eta\eta'}{2N} \cdot X_2 \\ -X_2' \cdot \frac{\eta\eta'}{2N} \cdot X_1 & X_2'(I - \frac{\eta\eta'}{2N})X_2 \end{bmatrix}^{-1}.$$