# Web Appendix W1Detailed Description of Initiatives of Policy Makers and the Online Advertising Industry to Restrict the Lifetime of Tracking Technologies

Online tracking in Europe is mainly governed by the ePrivacy Directive and the General Data Protection Regulation (GDPR). To date, these European laws have not restricted the lifetime of cookies or other tracking technologies. Still, some EU member states, such as Italy, France, and Spain, have already enacted such restrictions. The European Union is considering such restrictions on top of the current privacy laws, for example, within the upcoming ePrivacy Regulation (Council of the European Union 2021). This void in European regulation has led to various proposals on cookie lifetime and data retention periods from national data protection agencies and the online advertising industry, with unclear economic consequences for the online advertising industry.

 As early as 2010, the Article 29 Data Protection Working Party (2010), now superseded by the European Data Protection Board (<https://edpb.europa.eu/edpb_en>), published an opinion on online behavioral advertising with some clarifications on how to interpret the ePrivacy Directive. The Working Party (p. 6) acknowledges that “cookies have different life spans. This lifespan might or might not be extended in the future upon further visits to the same site (this is a design decision by the programmer). ‘Persistent cookies’ either have a precise expiry date far in the future or until they are manually deleted.”

 Further, the Working Party (p. 16) outlines that “users’ acceptance of a cookie could be understood to be valid not only for the sending of the cookie but also for subsequent collection of data arising from such a cookie. In other words, the consent obtained to place the cookie and use the information to send targeted advertising would cover subsequent 'readings' of the cookie every time the user visits a website partner of the ad network provider which initially placed the cookie. However, taking into account that i) this practice would mean that individuals accept to be monitored ‘once for ever,’ and, ii) individuals might simply ‘forget’ that, for example, a year ago, they agreed to be monitored; the Working Party considers that some safeguards should be implemented.”

Specifically, the Working Party (p. 16) requires the online advertising industry “to limit the scope of the consent in terms of time. Consent to be monitored should not be ‘for ever’, but it should be valid for a limited period of time, for example, one year. After this period, ad network providers would need to obtain new consent. This setting could be achieved if cookies had a limited lifespan after they have been placed in the user's terminal equipment (and the expiry date should not be prolonged).” Article 29 Data Protection Working Party (2010; p. 24) adds, “Ad network providers should implement retention policies which ensure that information collected each time that a cookie is read is automatically deleted after a justified period of time (necessary for the purposes of the processing).” Article 29 Data Protection Working Party (2011; p. 7) specifies the following: “Furthermore, the collection and processing of data for behavioral advertising purposes must be kept to a minimum.” Further, the Working Party (2011; pp. 7–8) criticizes that industry self-regulation efforts so far do “not contain any provisions on the amount of data collected and the retention period(s) for the specific purposes.… Such an initiative should at least address the period in which consent can be considered valid, and after which data shall then be deleted.”

 In 2013, Article 29 Data Protection Working Party (p. 3) further outlined the conditions for how to obtain valid consent from users: “Necessary information would be the purpose(s) of the cookies and, if relevant, an indication of possible cookies from third parties or third party access to data collected by the cookies on the website. Information such as the retention period (i.e., the cookie expiry date), typical values, details of third-party cookies and other technical information should also be included to fully inform users.”

 In 2015, Article 29 Working Party (p. 2) commissioned a cookie sweep of 478 websites in the e-commerce, media and public sectors across eight member states in the EU and “highlighted areas for improvement including a few cookies with duration periods of up nearly 8,000 years. This large value is in contrast to an average duration of 1 to 2 years. The sweep, however, also showed that 70% of the 16,555 cookies recorded were third-party cookies. Only 25 third-party domains set more than half of the third-party cookies.” Article 29 Working Party (2015, p. 19) continues, “Some first[-] and third[-]party cookies appear to have an extremely long duration. Cookies with an expiry set to 31/12/9999 23:59 (the maximum possible value) could be regarded as not having a reasoned retention schedule defined. Excluding cookies with a long duration (greater than 100 years), the average duration was between 1 to 2 years. This could be a useful starting point for a discussion regarding an acceptable maximum duration, although the purpose of the cookie will also need to be taken into account.“

 Again, neither the ePrivacy Directive nor the GDPR contains an explicit time limit for the validity of consent, the appropriate maximum cookie lifetime and the retention period for the associated data. To fill this void in the existing legal framework, Article 29 Data Protection Working Party (2010, p. 16) recommends an overall cookie lifetime of not more than one year. The European Parliament has circulated draft versions of the new ePrivacy Regulation (ePR) with cookie lifetime requirements from six months to no longer than 12 months (European Union 2017a; 2017b; Council of the European Union 2021). The French data protection agency has already implemented a maximum cookie lifetime of six months in France (Commission Nationale de l'Informatique et des Libertés 2020)[[1]](#footnote-2). In contrast, the Italian data protection agency requires deleting cookies in Italy after 12 months (Garante per la Protezione dei Dati Personali 2015; p. 17), and the Spanish Agencia Española de Protección de Datos (2020; p. 29) requires deleting cookies after 24 months. In Germany and the UK, data protection authorities acknowledge cookie lifespan restrictions as necessary and advocate relatively shorter lifespans but do not specify how long the useful life of a cookie should be (Voisin et al. 2021).

 In addition to the planned and already implemented initiatives of the various data protection authorities and regulatory bodies across Europe, various online advertising industry stakeholders have put forward their guidelines to restrict the lifetime of cookie tracking. For example, Tradelab, a sizable programmatic media buyer in Europe, has voluntarily committed to a maximum cookie lifetime of 13 non-rolling months and deletes the associated data after 12 months (Tradelab 2019). Google anonymizes advertising data by removing part of the IP address after nine months and their first-party cookie information after 18 months (Google 2022). Facebook keeps its first-party cookies for 24 months (Cook 2017). At the same time, prominent online advertising industry stakeholders such as Apple, Microsoft, Mozilla, and Google have either implemented or plan to completely ban third-party cookies (possibly also for the benefit of their first-party cookies) (Barker and Murgia 2020).

 Furthermore, various cookie and online consent solutions claim to comply with the EU ePrivacy Directive and the GDPR, such as Cookiebot (2022), OneTrust (2022), Curac-Dahl and Juszczyński (2021), and Koch (2022). They allow a maximum cookie lifetime of 12 months (see also Sanchez-Rola et al. 2019).

# Web Appendix W2Exemplary Observed Maximum Cookie Lifetimes from Selected Domains

TABLE W2.1
EXEMPLARY OBSERVED MAXIMUM COOKIE LIFETIMES
FROM SELECTED DOMAINS

|  |  |  |  |
| --- | --- | --- | --- |
| Cookie Domain | Cookie Name(given by owner of cookie) | Expiration Date | Maximum Cookie Lifetime(in days) |
| amazon.com | \_utma | March 13, 2025 | 730 |
| aws-at-main | February 9, 2043 | 7,272 |
| aws-priv | March 14, 2026 | 1,096 |
| google.com | anid | April 7, 2024 | 390 |
| apisid | March 13, 2025 | 730 |
| consent | October 19, 2041 | 6,794 |
| bing.com | muid | April 7, 2024 | 390 |
| srchid | April 7, 2024 | 390 |
| srchuid | April 7, 2024 | 390 |
| facebook.com | c\_user | March 13, 2024 | 365 |
| datr | March 13, 2025 | 730 |
| dbln | March 13, 2025 | 730 |
| twitter.com | \_ga | March 13, 2025 | 730 |
| ads\_prefs | March 12, 2028 | 1,825 |
| dnt | March 12, 2028 | 1,825 |
| Notes: Mean maximum cookie lifetime = 1,626 days (min = 365 days, max = 7,272 days). Cookies are generated from a single visit to the respective domains after deleting all old cookies and the entire browsing history on March 14, 2023, using Google Chrome (chrome://settings/siteData).  |

# Web Appendix W3Survival Model Validation

We use a third sample to validate our approach to determine the residual cookie lifetime for potentially censored cookies. To construct the sample, we focus on a cohort of new cookies to avoid any potential left-censoring and to be able to observe the actual cookie birthday in our data. We consider a cookie to be a member of this newborn cohort of cookies if the cookie is active for the first time in our focal sampling week but was not active at any time in the ten weeks before the beginning of our observation period. Thus, it is very likely that all cookies we observe for the first time in calendar week 20 (i.e., in the 11th week of our data set) are newborn cookies. This selection leaves us with 30,952 cookies in our subsample of newborn cookies. Only 18 cookies (.058%) are potentially right-censored (i.e., active within the last seven days of our observation period), but no cookies are left-censored. Thus, we have a maximum of 797 observable days for the newborn cookie cohort.

 The third sample differs from the first and second samples as it includes only the newborn cookies in week 20 of our data. In contrast, samples 1 and 2 include cookies born on the respective sampling days and those born before our sampling days. It leads, for example, to a much larger observed cookie lifetime in samples 1 and 2 than in sample 3. This difference in composition of the different samples used in our study is, however, no reason for concern regarding the validation of our survival models used in our study. We use the newborn cohort of the third sample only for validation, not to obtain our main results.

 Our descriptive analysis of the subsample of newborn cookies, as summarized in Table W3.1, shows a mean cookie age on our sampling day of 0 days, as expected. We find a mean cookie lifetime of 35 days (1.2 months) and a median of 1 day. The observed cookies differ concerning the number of ad impressions served. They reach an average number of ad impressions of about 142 and a median of four ad impressions. We calculate that cookie values per day in our data set reach a mean value of €.006 and a median of €.001. Finally, the observed mean price per 1,000 ad impressions paid by the purchasing advertiser is €.807 CPM, and the median price is €.709 CPM. Concerning the observed cookie lifetime value, we find a mean lifetime value of €.095 and a median of €.004.

 We split the newborn subsample into a training and test data set to validate our approach to determine residual cookie lifetime. We use the first 500 days (63%) as our estimation sample and estimate a parametric Weibull, Lognormal, and Generalized Gamma model. We report the model fit measures in Table W3.2 Panel 1. The Generalized Gamma model fits best, but the differences between all models are relatively small.

 We use the remaining 297 days (37%) of our data as our holdout sample and report the respective validation measures in Panel 2 of Table W3.2. We follow Meeker and Escobar’s (1998) suggestion to obtain the mean and median residual lifetime per cookie (RLT) as outlined in the main part of the paper. The Weibull model best predicts the observed mean cookie lifetime in our newborn sample of 34.100 [95% confidence interval (95% CI): 32.998; 35.203] with an estimated uncensored mean cookie lifetime of 35.380 [95% CI: 34.194; 36.566]. The Weibull model also shows the best validation measures with an R-squared between the observed and predicted values of cookie lifetime of .974, a mean absolute error (MAE) of 1.958, a root mean squared error (RMSE) of 18.221 and a mean absolute percentage error (MAPE) of .003.

 Taken together, the parametric Weibull model provides the best fit. Figure W3.1 illustrates this conclusion by plotting the observed cookie lifetime and the Weibull model fit. Therefore, we subsequently use the parametric Weibull model to determine the uncensored mean cookie lifetime in samples 1 and 2 in the main part of the manuscript.

TABLE W3.1
SAMPLE 3: SUMMARY STATISTICS PER COOKIE – NEWBORN COHORT IN CALENDAR WEEK 20 OF 2014
(N = 30,952)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Category | Variable | Quantiles | Mean | SD |
| Min. | 25% | 50% | 75% | Max. |
| CookieLifetime Units | Observed Cookie Age on Sampling Day (in days)a | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Observed Cookie Lifetime (in days)ab | 1 | 1 | 1 | 9 | 791 | 35 | 99 |
| Observed Number of Ad Impressions | 1 | 1 | 4 | 18 | 87,313 | 142 | 1,497 |
| Cookie Valueper LifetimeUnit | Observed Cookie Value per Day (in €) | .000 | .000 | .001 | .005 | 5.000 | .006 | .039 |
| Observed Cookie Value per Ad Impression (in €, CPM) | .000 | 411 | .709 | 1.096 | 27.997 | .807 | .853 |
| CookieLifetime Value | Observed Cookie Lifetime Value (in €) | .000 | .001 | .004 | .018 | 56.059 | .095 | .908 |
| Note: a Rounded to the next full day. b Using the data of sample 3, we generate a subsample of new cookies born in calendar week 20 in 2014. Our observation window starts in calendar week 10 in 2014, so we can use weeks 1–9 to check that the cookie was not observed before. This long period makes it very likely that all cookies observed for the first time in calendar week 20 are newborn cookies, so no cookies are left-censored. We have a maximum of 797 observable days for the newborn cookie cohort. Of these newborn cookies, only 18 cookies (.058% of all cookies in the subsample) are potentially right-censored (i.e., lived for more than 797 days).  |

TABLE W3.2
SAMPLE 3: SURVIVAL MODEL VALIDATION – NEWBORN COHORT IN CALENDAR WEEK 20 OF 2014 (N = 30,952)

|  |
| --- |
| *Panel 1: Model Fit Measures* |
| Model | Shape Parameter[95% CI] | SE | Scale Parameter[95%- CI] | SE | LL | AIC | BIC |
| Weibull | .823[.808; .837] | .007 | 114.302[111.136; 117.558] | 1.638 | −44,898.810 | 89,801.620 | 89,815.650 |
| Lognormal | 4.120[4.093; 4.148] | .014 | 1.282[1.262; 1.303] | .011 | −44,195.650 | 88,395.290 | 88,409.320 |
| Generalized Gamma | 3.664[3.606; 3.722] | .030 | 1.190[1.164; 1.216] | .013 | −44,039.840 | 88,085.670 | 88,106.710 |
| *Panel 2: Validation Measures* |
| Model | Observed MeanCookie Lifetime[95% CI] | Uncensored MeanCookie Lifetime (Mean RLT)a[95%-CI] | R2 | MAE | RMSE | MAPE |
| Weibull | 34.100[32.998; 35.203] | 35.380[34.194; 36.566] | .974 | 1.958 | 18.221 | .003 |
| Lognormal | 40.197[38.643; 41.752] | .925 | 6.097 | 52.314 | .010 |
| Generalized Gamma | 106.883[99.472; 114.294] | .659 | 72.783 | 592.174 | .117 |
| Model | Uncensored Mean Cookie Lifetime (Median RLT)b[95%-CI] | R2 | MAE | RMSE | MAPE |
| Weibull | 108.679[106.424; 110.933] | .896 | 74.578 | 135.616 | 6.867 |
| Lognormal | 70.230[68.655; 71.805] | .939 | 36.130 | 63.039 | 3.891 |
| Generalized Gamma | 53.027[50.371; 55.683] | .827 | 18.926 | 155.187 | .030 |
| Notes: To avoid cookies with very short lifetimes impacting our results too strongly, we only consider cookies with an observed cookie lifetime of seven or more days to predict residual cookie lifetime. Under the lognormal model, the shape parameter equals the mean, and the scale parameter equals the standard deviation. The generalized gamma model has an additional parameter kappa, which is k = −.742 [−.827; −.657], SE = .043. a We determine the uncensored mean cookie lifetime using the predicted mean residual lifetime (RLT). b We determine the uncensored mean cookie lifetime using the predicted median residual lifetime (RLT). SE: standard error; LL: log-likelihood; AIC: Akaike information criterion; BIC: Bayesian information criterion; RLT: residual lifetime; MAE: mean absolute error; RMSE: root mean squared error; MAPE: mean absolute percentage error. |

FIGURE W3.1
SAMPLE 3: OBSERVED COOKIE LIFETIME AND WEIBULL MODEL FIT FOR NEWBORN COHORT IN CALENDAR WEEK 20 OF 2014 (N = 30,952)



 Another cause for concern is the 7-day threshold to determine whether we observe the birth or death of a cookie. In doing so, we are prone to two types of errors: First, we extend the observed lifetime of a cookie, although the cookie was actually deleted. Second, we do not extend the observed lifetime of a cookie, although the cookie lived longer than the observed lifetime.

 If we choose a rather short threshold (i.e., seven days as in our main analysis), we are more likely to not extend the observed lifetime of a cookie, although the cookie lived longer. Alternatively, suppose we choose a rather long threshold (i.e., 28 days, as in our following robustness analysis). In that case, we are more likely to overestimate the lifetime of a cookie, although the cookie was deleted.

 Our main specification in the paper uses a 7-day threshold to account for potential censoring of the observed cookie lifetime. Specifically, we use a survival model to compute the predicted residual mean lifetime for those cookies that received ad impressions (and thus were active) within the first and last seven days of our observation period. Based on this cut-off criterion, 13.123% of all cookies are (potentially) censored: 5.572% are left-censored (i.e., we are only able to observe cookie death), 5.271% are right-censored (i.e., we are only able to observe cookie birth), and 2.280% are both right- and left-censored (i.e., we observe neither cookie birth nor cookie death).

 For robustness, we extend the threshold for user inactivity and compute the predicted residual mean lifetime for those cookies that received ad impressions (and thus were active) within the first and last 28 days of our observation period. Based on this cut-off criterion, 17.656% of all cookies are (potentially) censored. 7.015% are left-censored, 6.889% are right-censored, and 3.752% are right- and left-censored.

 Using the 7-day (28-day) threshold, we obtain an average uncensored lifetime of a cookie of 279 (328) days and an uncensored lifetime value of a cookie of €2.522 (€2.958). That is, the uncensored lifetime of a cookie under a 28-day threshold is 17.563% larger than under the 7-day threshold, and the uncensored lifetime value of a cookie is 17.288% larger under the 28-day threshold than under the 7-day threshold (see Table W3.3).

 Although we predict a larger lifetime and value of the cookies under the 28-day threshold, we only find a slightly larger economic loss of the various cookie lifetime restrictions (see Table W3.4). For example, under the 7-day (28-day) restriction, we find a 13.916% average %-loss (versus 15.888%) for the 30-day restriction. We also find a 5.432% average %-loss for the 720-day restriction (versus 7.403%). The differences in the average %-loss between the 7-day and the 28-day threshold are statistically significant, as the 95%-confidence intervals do not overlap.

 In summary, our choice of a rather short threshold of seven days to determine censored observations in our data does not lead us to overestimate the economic loss of various cookie lifetime restrictions. Instead, our results provide a lower bound for economic loss. Increasing the threshold to 28 days only slightly increases our loss estimates, providing additional confidence in our results' robustness.

TABLE W3.3
ROBUSTNESS TO USER INACTIVITY THRESHOLD (N = 54,127)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Threshold for User Inactivity | Variable | Quantiles | Mean | SD |
| Min. | 25% | 50% | 75% | Max. |
| - | Observed (potentially censored) Lifetime of Cookie (in days) a | 1 | 1 | 68 | 416 | 867 | 216 | 268 |
| Observed (potentially censored) Lifetime Value of Cookie (in €) | .000 | .003 | .022 | .498 | 331.048 | 1.428 | 5.234 |
| 7 days | Uncensored Lifetime of Cookie (in days) ab | 1 | 1 | 68 | 416 | 1,351 | 279 | 396 |
| Uncensored Lifetime Value of Cookie (in €) c | .000 | .003 | .022 | .610 | 449.403 | 2.522 | 10.603 |
| 28 days | Uncensored Lifetime of Cookie (in days) ad | 1 | 1 | 68 | 420 | 1,516 | 328 | 481 |
| Uncensored Lifetime Value of Cookie (in €) e | .000 | .003 | .022 | .678 | 555.438 | 2.958 | 12.765 |
| Note: a Rounded to the next full day. b We use a Weibull model to determine the expected residual lifetime for 13.123% of the cookies with potentially censored cookie lifetime and determine censored cookies based on a threshold of 7 days of inactivity from the beginning and the end of our observation period. The adjusted average predicted cookie lifetime of 279 days is 29.167% larger than the average observed cookie lifetime of 216 days in the data (i.e., sample 1). c We determine the uncensored cookie lifetime value using the regression from Equation 1 (i.e., model 2 in Table 7). d  We use a Weibull model to determine the expected residual lifetime for 17.656% of the cookies with potentially censored cookie lifetime. We determine the censored cookies based on a threshold of 28 days of inactivity from the beginning and end of our observation period. The adjusted average predicted cookie lifetime of 328 days is 51.851% larger than the average observed cookie lifetime of 216 days in the data (i.e., sample 1) and 17.563% larger than the adjusted average predicted cookie lifetime of 279 days.  |

TABLE W3.4
ECONOMIC LOSS OF VARIOUS COOKIE LIFETIME RESTRICTIONS

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Cookie Lifetime Restriction* | *Fulfillment of Condition Ia* | *Fulfillment of Conditions I & IIb* | *Fulfillment of Conditions I & IIIc* | *Economic Loss per Cookie that Fulfills Conditions I & II* | *Economic Loss per Cookie that Fulfills Conditions I & III* | *Economic Loss per Cookie* |
| No.Cookies | % of all Cookies | % Cond.IId | % Cond.IIIe | NoCookies | % of all Cookies | NoCookies | % of all Cookies) | Average LVC | Average Absolute[95%-CI] | Average % Loss[95%-CI] | Average LVC | Average Absolute[95%-CI] | Average % Loss[95%-CI] | Average LVC | Average Absolute[95%-CI] | Average % Loss[95%-CI] |
| *Panel 1: Simulation results based on Sample 1 (N = 54,127): Regression of average price per ad impression per day on day count + additional covariates (Model 2), 7 day threshold for user inactivity* |
| 30 days | 29,647 | 54.773% | 22.410% | 13.769% | 6,644 | 12.275% | 4.082 | 7.542% | €10.966 | €4.135[4.074; 4.196] | 37.707%[37.149; 38.265] | €4.906 | −€2.082[−2.122; −2.041] | −42.434%[−43.260; −41.609] | €2.522 | €.351[.340; .361] | 13.916%[13.480; 14.312] |
| 360 days | 15,329 | 28.320% | 28.234% | 16.413% | 4,328 | 7.996% | 2,516 | 4.648% | €15.332 | €3.785[3.714; 3.856] | 24.686%[24.226; 25.146] | €7.025 | −€1.883[−1.949; −1.817] | −26.800%[−27.737; −25.864] | €2.522 | €.215[.206; .224] | 8.524%[8.167; 8.881] |
| 720 days | 7,953 | 14.693% | 35.986% | 17.805% | 2,862 | 5.288% | 1,416 | 2.616% | €20.401 | €3.796[3.723; 3.869] | 18.607%[18.248; 18.966] | €9.863 | −€2.454[−2.569; −2.339] | −24.877%[−26.044; −23.711] | €2.522 | €.137[.130; .143] | 5.432%[5.154; 5.669] |
| *Panel 2: Simulation results based on Sample 1 (N = 54,127): Regression of average price per ad impression per day on day count + additional covariates (Model 2), 28 day threshold for user inactivity* |
| 30 days | 29,647 | 54.773% | 22.410% | 13.769% | 6,644 | 12.275% | 4.082 | 7.542% | €13.413 | €5.362[5.286; 5.438] | 39.976%[39.407; 40.546] | €5.236 | −€2.495[−2.543; −2.447] | −47.651%[−48.571; −46.731] | €2.958 | €.470[.457; .483] | 15.888%[15.449; 16.328] |
| 360 days | 15,329 | 28.320% | 28.234% | 16.413% | 4,328 | 7.996% | 2,516 | 4.648% | €19.089 | €5.607[5.509; 5.705] | 29.374%[28.860; 29.887] | €7.561 | −€2.796[−2.888; −2.703] | −36.974%[−38.195; −35.753] | €2.958 | €.318[.306; .331]. | 10.750%[10.344; 11.189] |
| 720 days | 9,777 | 18.063% | 33.477% | 16.948% | 3,273 | 6.047% | 1,657 | 3.061% | €23.428 | €5.435[5.350; 5.21] | 23.201%[22.837; 23.564] | €9.566 | −€3.578[−3.710; −3.446] | −37.401%[−38.779; −36.024] | €2.958 | €.219[.210;.228] | 7.403%[7.099; 7.708] |
| Notes: a Condition I refers to the number of cookies with a cookie lifetime larger than the cookie lifetime restriction. Condition II refers to the number of cookies that increase their values per day. Condition III refers to the number of cookies that decrease their values per day. b Conditions I and II refer to the number of cookies that fulfill condition I and increase their values per day. c Conditions I and III refer to the number of cookies that fulfill condition I and decrease their values per day. d Share of those cookies that also fulfill condition II (P(Cond.II | Cond. I) = P(Cond. I & II) / P(Cond. I). e Share of those cookies that also fulfill condition III (P(Cond.III | Cond. I) = P(Cond. I & III) / P(Cond. I).Reading example: 29,647 cookies (i.e., 54.773% of all cookies) fulfill condition I (i.e., have a cookie lifetime larger than the imposed cookie lifetime restriction of 30 days), and 22.410% of these cookies fulfill condition II and 13.769% condition III. 6,644 cookies (i.e., 12.275% of all cookies) fulfill conditions I and II (i.e., increase in value per day). 4,082 cookies (i.e., 7.542% of all cookies) fulfill conditions I and III (i.e., decrease in value per day). The average cookie that fulfills conditions I and II has an average cookie lifetime value (LVC) of €10.966 and loses under a 30-day lifetime restriction on average €4.135 (or 37.707% of the total average LVC). The average cookie that fulfills conditions I and III has an average LVC of €4.906 and loses under a 30-day lifetime restriction on average −$.2.082 (i.e., −42.434% of the total average LVC). The average cookie in our first sample has an LVC of $2.522 and loses, on average, a value of €.351 (or 13.916%) under a 30-day cookie lifetime restriction policy. |

# Web Appendix W4Validation of Regressions

We run four other regression models to validate our regression results to determine the incremental cookie value per lifetime unit (Step 4 in Figure 3 in the main part of the manuscript).

 In models 1 and 2, we use the average price per ad impression per day as the dependent variable. As an independent variable, we use the day count, which captures the incremental value of time. In model 1, we only use the day count variable as a covariate, while in model 2, we consider ad inventory characteristics as additional covariates. These ad inventory characteristics control for media type, which captures the share of video ads over regular display ads per day; fold position, which captures the share of ads displayed above the fold per day; and the share of retargeted ad impressions per day.

 In models 3 and 4, we use ln(average price per ad impression per day) as the dependent variable. As an independent variable, we use day count without additional covariates (model 3) and day count and ad inventory characteristics as additional covariates (model 4).

 Finally, in model 5, we use the average price per ad impression per day as the dependent variable. As an independent variable, we use day count, day count2 and ad inventory characteristics as additional covariates.

 Our main criterion for model selection is prediction quality (see Panel 1 in Table W4.1). Therefore, we calculate the prediction quality measures R-squared, mean average error (MAE), and the root mean squared error (RMSE) using the first 80% of consecutive observations of each cookie to train the model and the last 20% observations to test the model. Finally, we calculate the mean absolute percentage error (MAPE), corresponding to the in-sample absolute difference between the cookie's observed lifetime value and the cookie's predicted lifetime value divided by the cookie's observed lifetime value. Our preferred specification, the linear model with ad inventory characteristics as additional covariates (model 2), performs best on average, for example, with regard to MAPE. We, therefore, choose model 2 as our primary model in the main manuscript, which we use to determine the incremental cookie value per lifetime unit.

 As model fit measures, we determine the R-squared, the Akaike information criterion (AIC), and the Bayesian information criterion (BIC) using the entire available data set (see Panel II in Table W4.1). With regard to model fit, the log-linear model (model 4) and the quadratic model (model 5) show a slight improvement in R-squared over model 2, but have a worse (model 4) or comparable fit (model 5) when comparing the Akaike information criterion (AIC), and the Bayesian information criterion (BIC).

 We repeat the abovementioned analysis regarding our prediction quality and model fit measures for our second sample and obtain similar results. The detailed results are available from the authors upon request.

TABLE W4.1
ORDINARY LEAST SQUARES PREDICTION QUALITY AND MODEL FIT MEASURES

| *Model Number* | *Dependent Variable* | *Independent Variable(s)* | *Measure* | *Quantiles* | *Mean* | *SD* |
| --- | --- | --- | --- | --- | --- | --- |
| *Min* | *25%* | *50%* | *75%* | *Max* |
| *Panel I: Prediction Quality Measuresa* |
| 1 | Averageprice per ad impression per day | Day count | R2 | .000 | .029 | .117 | .367 | 1.000 | .261 | .318 |
| MAE | .000 | .247 | .390 | .624 | 36.410 | .526 | .629 |
| RMSE | .000 | .311 | .501 | .816 | 44.533 | .683 | .817 |
| MAPE | .000 | .046 | .108 | .223 | 13.468 | .194 | .348 |
| 2 | Averageprice per ad impression per day | Day countad inventory | R2 | .000 | .025 | .116 | .425 | 1.000 | .274 | .330 |
| MAE | .000 | .246 | .395 | .633 | 36.410 | .535 | .670 |
| RMSE | .000 | .312 | .510 | .833 | 44.533 | .700 | .857 |
| MAPE | .000 | .043 | .100 | .204 | 11.222 | .175 | .304 |
| 3 | ln (Averageprice per ad impression per day) | Day count | R2 | .000 | .032 | .128 | .388 | 1.000 | .269 | .318 |
| MAE | .001 | .357 | .485 | .671 | 16.937 | .592 | .502 |
| RMSE | .001 | .446 | .600 | .809 | 17.018 | .706 | .541 |
| MAPE | .000 | .061 | .138 | .263 | 10.814 | .202 | .246 |
| 4 | ln (Averageprice per ad impression per day) | Day countad inventory | R2 | .000 | .031 | .127 | .428 | 1.000 | .279 | .327 |
| MAE | .001 | .355 | .484 | .670 | 11.991 | .594 | .509 |
| RMSE | .001 | .444 | .598 | .810 | 16.177 | .711 | .560 |
| MAPE | .000 | .052 | .120 | .240 | 41.495 | .185 | .338 |
| 5 | Averageprice per ad impression per day | Day countday count2ad inventory | R2 | .000 | .028 | .117 | .399 | 1.000 | .269 | .325 |
| MAE | .000 | .290 | .492 | .867 | 280.137 | .890 | 3.493 |
| RMSE | .000 | .371 | .635 | 1.115 | 384.279 | 1.113 | 4.358 |
| MAPE | .000 | .043 | .102 | .212 | 12.305 | .181 | .315 |
| Panel II: Model Fit Measuresb |
| 1 | Averageprice per ad impression per day | Day count | R2 | .000 | .006 | .028 | .088 | .953 | .073 | .113 |
| AIC | −1,329.290 | 29.730 | 106.460 | 343.810 | 4,953.150 | 236.820 | 366.717 |
| BIC | −1,321.630 | 33.180 | 112.780 | 353.910 | 4,967.200 | 243.970 | 369.104 |
| 2 | Averageprice per ad impression per day | Day countad inventory | R2 | .000 | .033 | .088 | .205 | .998 | .151 | .168 |
| AIC | −1,329.290 | 27.030 | 101.780 | 335.430 | 4,951.780 | 230.750 | 366.594 |
| BIC | −1,321.63 | 31.320 | 110.350 | 349.160 | 4,970.52 | 240.380 | 370.042 |
| 3 | ln (Averageprice per ad impression per day) | Day count | R2 | .000 | .009 | .039 | .120 | .963 | .093 | .135 |
| AIC | −472.400 | 53.900 | 149.500 | 396.700 | 2201.300 | 266.300 | 287.872 |
| BIC | −464.750 | 57.830 | 156.650 | 406.960 | 2215.160 | 273.420 | 287.872 |
| 4 | ln (Averageprice per ad impression per day) | Day countad inventory | R2 | .000 | .052 | .122 | .246 | .989 | .176 | .169 |
| AIC | −472.410 | 52.080 | 144.760 | 386.140 | 2149.510 | 258.930 | 281.427 |
| BIC | −464.750 | 57.110 | 154.300 | 399.880 | 2168.050 | 268.570 | 285.693 |
| 5 | Averageprice per ad impression per day | Day countday count2ad inventory | R2 | .000 | .060 | .131 | .269 | .998 | .196 | .187 |
| AIC | -1.329.450 | 26.570 | 100.580 | 330.400 | 4,952.970 | 227.510 | 364.444 |
| BIC | -1.319.230 | 31.970 | 111.380 | 347.230 | 4,976.400 | 239.530 | 368.654 |
| Notes: This table reports estimates of cookie-specific regressions per reported model conditional on having at least ten observations per cookie. The resulting number of observations is 28,788. Ad inventory characteristics include media type, which captures the share of video ads over regular display ads per day; fold position, which captures the share of ads displayed above the fold per day; and the share of retargeted ad impressions per day. a The prediction quality measures R2, MAE, and RMSE are obtained by training the model on the first 80% of consecutive observations per cookie and testing the model on the last 20% of consecutive observations per cookie. MAPE corresponds to the in-sample absolute difference between the observed lifetime value of the cookie and the predicted lifetime value of the cookie divided by the observed lifetime value of the cookie using the full data set. AIC: Akaike information criterion; BIC: Bayesian information criterion; MAE: mean absolute error; RMSE: root mean squared error; MAPE: mean absolute percentage error. b Model fit measures are calculated on the full data set. |

# Web Appendix W5Simulation Study

In this Web Appendix, we outline how our simulation accommodates that (1) the predicted lifetime is longer than the observed lifetime, (2) cookies are not active every day, and (3) eliminates differences in the daily number of impressions per cookie across time.

## Consideration of Difference Between Predicted and Observed Lifetime of a Cookie

Due to censoring, a cookie’s actual lifetime (herein referred to as the uncensored lifetime of a cookie) can be larger than its observed (potentially censored) lifetime; thus, we must predict the price per ad impression for each day of the predicted residual lifetime beyond the observed lifetime. We use Cookie A to illustrate how we proceed if the observed lifetime is only 15 days but the uncensored lifetime is 22 days (as is the case for Cookie A in the main manuscript). In this case, the cookie generates a value of $.09 on day 1, $.10 on day 2, $.11 on day 3, and so forth, until $.23 on day 15. The observed lifetime value of this cookie (LVC)is thus the sum of the cookie values per day across the 15 days ($.09 + $.10 + $.11 +… + $.23 = $2.40). We then estimate our regression. The estimated time parameter outlines that the value per day increases by $.01. We use this information to predict the value per day for the remaining seven days (i.e., $.24 on day 16, $.25 on day 17, and so forth, to $.30 on day 22). The uncensored LVCis thus the sum of the cookie values per day across the 15 observed days (days 1–15: $2.40) and the 7 additional days (days 16–22: $1.89), thus $4.29.

## Consideration of the Difference Between the Number of Active and Observed Days

So far, we have looked at cookies—more precisely, the users behind each cookie who were active each day (i.e., received an ad impression each day). Yet, cookies are also inactive on some days (i.e., they receive no ad impressions). We use Cookie A again to illustrate how we consider inactivity in our simulation study. Cookie A again has an observed lifetime of 15 days, and its uncensored lifetime is 22 days. Yet, Cookie A was inactive on day 7, that is, on one of the 15 days (= 1 / 15 = 6.66%). Thus, the share of observed active days per observed days of Cookie A is 93.33%.

 We again estimate the regression and use the time parameter to predict the price of the ad impression on the inactive day (day 7) and the additional 7 days (days 16–22). The sum of those values is again $4.29. However, we now multiply the total value by the share of active days: 93.33% × $4.29 = $4.00. We also multiply the total value with the restriction ($2.89) with the share of active days, resulting in $2.70. As a result, LVC is reduced by $1.30, from $4.00 to $2.70 (−32%).

## Consideration of a Trend in the Number of Ad Impressions per Day

So far, we have considered an equal number of ad impressions per day, namely one, such that the average price of an ad impression was equal to the revenue per day. However, revenue is the product of the number of ad impressions per day times the average daily price per ad impression; therefore, we must account for the difference between revenue and price per ad impression. Only changes in the price per ad impression reflect that an increase in information about the user (as made available by the cookie) reflects the cookie’s value.

 It is possible that the cookie changed activity over time; for example, it could become more active because the user behind the cookie learned how to use the internet better. These changes do not reflect the value of the cookie. Therefore, we need to isolate the changes in revenue per day from those changes due to changes in the average price, not those due to changes in the number of ad impressions.

 We again use Cookie A to illustrate how we isolate those changes in our simulation study. Cookie A again has an observed lifetime of 15 days, its uncensored lifetime is 22 days, and it was inactive on day 7, that is, on one of the 15 days (= 1 / 15 = 6.66%). Thus, the share of observed active days at all observed days of Cookie A is 93.33%. Yet, daily ad impressions no longer remain constant but increase by 1 per day. So, we have 1 ad impression on day 1, 2 ad impressions on day 2, and so forth, until 15 ad impressions on day 15. The revenue of those 15 days is now $20.95.

 Now, we run two regressions — one for the daily price (as we did before) and one for the daily ad impressions. The result for the “price regression” is (as before) price per impression and day = $.08 + $.1 × day. The result for the “quantity regression” is impressions per day = 0 + 1 × day. Thus, the time parameter in this regression represents the daily increase in ad impressions. We eliminate these changes in the number of ad impressions by replacing the daily number of impressions with the daily number of impressions predicted by the regression for the average value of the independent variable, which is 8.07 (= 0 + 1 × 8.07, as we ignore the inactive day 7). This value also equals the average daily number of impressions in the first 15 days (actually 14 days, as we ignore the inactive day 7), which is 8.07 (=113 / 14).

 We then proceed as outlined previously. We calculate the revenue for the inactive day 7 ($1.21 = 8.07 × $0.15) and days 16–22 (e.g., for day 16: $1.94 = 8.07 × $.24). We multiply the sum of those values with the share of active days ($34.05 = 93.33% × $36.48), representing Cookie A’s value without lifetime restrictions. We proceed similarly for the case of the lifetime restriction, yielding a value of $22.75 and thus a loss of $11.30 (−33.19%).

 In the Supplemental Material to this manuscript, we provide a spreadsheet that outlines how we determine the economic consequences of cookie lifetime restrictions in our numerical example.

# Web Appendix W6Descriptive Regressions

TABLE W6.1
REGRESSION RESULTS OF COVARIATES PER COOKIE
ON CONSTANT AND TIME PARAMETER

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | Dep. Variable:Constant | p-Value | Dep. Variable:Time Parametera | p-Value |
| Country | National | .441 | .000\*\*\* | .212 | .837 |
| International | .308 | .004\*\* | −.171 | .868 |
| DeviceType | Desktop | .168 | .000\*\*\* | .158 | .303 |
| Mobile | .237 | .000\*\*\* | .325 | .160 |
| Operating System | Android | −.242 | .000\*\*\* | −.289 | .128 |
| BlackBerry | −.327 | .000\*\*\* | −.557 | .004\*\* |
| Chrome | .120 | .000\*\*\* | −.429 | .050\* |
| iOS | −.071 | .000\*\*\* | −.394 | .032\* |
| Linux | −.156 | .000\*\*\* | −.004 | .986 |
| PlayStation | .170 | .000\*\*\* | −.169 | .378 |
| Windows | −.040 | .000\*\*\* | .073 | .137 |
| Browser | Android | .232 | .000\*\*\* | −.140 | .722 |
| Chrome | .142 | .000\*\*\* | .012 | .974 |
| Firefox | .116 | .003\*\* | −.068 | .858 |
| InternetExplorer | .029 | .320 | .137 | .720 |
| iOS | −.120 | .006\*\* | −.282 | .450 |
| Opera | .189 | .000\*\*\* | .244 | .565 |
| Safari | .284 | .000\*\*\* | −.005 | .990 |
| Constant | — | −.212 | .061 | .076 | .944 |
| Adj. R2 | — | .036 | .001 |
| N | 98,527 |
| Notes: This table shows the results of two regressions using the pooled data of samples 1 and 2. p-values: \* p ≤ .05, \*\* p ≤ .01, \*\*\* p ≤ .001. The reference category for each categorical variable is “Unknown”. a We multiplied the parameter estimates by 1,000 for ease of readability. |

# Web Appendix W7Additional Results Sample 2

**Description of the Second Sample**

We repeat the analysis for our second sample, and Table W7.1 provides the descriptive statistics. The mean cookie age on our sampling day is 120 days (median: 28 days). We also find a mean observed cookie lifetime of 228 days (median: 94).

 We again use a survival model to account for potential censoring of the observed cookie lifetime. Our second sample includes 5.399% of (potentially) left-censored cookies (i.e., we can only observe cookie death) and 6.313% of right-censored cookies (i.e., we can only observe cookie birth). 2.550% are both right- and left-censored (i.e., we can observe neither cookie birth nor cookie death). Thus, we have 14.262% of censored cookies in our second sample. We again fit a parametric Weibull and parametric Lognormal survival model to the data (see Table W7.2).

 We select the Weibull model for further analysis because it fits the data better than the Lognormal model (Weibull: LL: −154,132.700, AIC: 308,269.500, BIC: 308,285.900 vs. Lognormal: LL: −154,630.200, AIC: 309,264.500, BIC: 309,280.900) (for validation of this approach, see Web Appendix W3). Again, we integrate the parametric Weibull distribution to find the residual cookie lifetime per cookie using the observed lifetime per cookie as the lower bound of the integral. The mean cookie lifetime adjusted for censoring is 298 days (median: 94 days).

 The cookies of our second sample were active on average on 81 of the observed days (median: 11 days), yielding an average share of observed active days at the observed days of 55.300% (median: 54.000%). The cookies also differ concerning the observed number of ad impressions served. They reach an average number of 2,369 ad impressions (median: 47) and an average number of ad impressions per observed day of 9.064 (median: 2.172).

 We calculate that cookie values per day in our second sample reach a mean value of €.005 (median: €.001). We also compute the average uncensored cookie value per day at €.006 (median: €.001). The mean value per 1,000 ad impressions paid by the purchasing advertiser is €.720 CPM (median: €.683 CPM).

 Concerning the observed cookie lifetime value, we find a mean lifetime value of €1.569 (median: €.030). The mean predicted lifetime value is €1.791 (median: €.029). We calculate a MAPE of .083 (median: .003). The mean predicted residual lifetime value amounts to €.960 (median: .000.). The uncensored cookie lifetime value yields different results, with a mean cookie lifetime value of €2.799 (median: €.029).

**Results of Regression Analysis**

Linear model without additional covariates.We summarize the results of our regression analysis in Table W7.3. In model 1, we find a positive incremental effect for 5,890 cookies. They represent 13.266% of all cookies in our sample. They received 46.669% of all ad impressions, with an average number of ad impressions per cookie of 8,336. Their mean uncensored cookie lifetime is 629 days, and the average uncensored cookie’s lifetime value is €10.123.

 We find a significant negative incremental effect for 4,376 cookies, representing 9.856% of all cookies in our sample. They received 18.942% of all ad impressions, with an average number of ad impressions per cookie of 4,554. Their mean cookie lifetime is 586 days, and the average uncensored cookie lifetime value is €4.950.

 For most cookies (34,134, i.e., 76.878%), the time parameter is insignificant, i.e., zero. Specifically, 10,196 cookies, or 22.964% of all cookies, have this “zero effect”. Yet, these cookies received 34.389% of all ad impressions, with an average number of ad impressions per cookie of 1,060. Their mean uncensored cookie lifetime is 203 days, and the average uncensored cookie lifetime value is €1.205.

 Linear model with additional covariates.In model 2, our preferred model specification (see Equation 1), we find a positive incremental effect for 6,106 cookies. They represent 13.752% of all cookies in our sample. They received 52.197% of all ad impressions, with an average number of ad impressions per cookie of 8,993. Their mean uncensored cookie lifetime is about 632 days, and the average uncensored cookie’s lifetime value is €10.756.

 We find a significant negative incremental effect for 4,097 cookies, representing 9.227% of all cookies in our sample. They received 16.245% of all ad impressions, with an average number of ad impressions per cookie of 4,172. Their mean uncensored cookie lifetime is 586 days, and the average uncensored cookie lifetime value is €4,761.

 The time parameter is insignificant for most cookies (34,197 or 77.020%). Specifically, 10,132 cookies, or 22.820% of all cookies, have a zero effect in the sample. Yet, these cookies received 31.558% of all ad impressions, with an average number of ad impressions per cookie of 971. Their mean uncensored cookie lifetime is 203 days, and the average uncensored cookie lifetime value is €1.143.

Table W7.1
SAMPLE 2: SUMMARY STATISTICS PER COOKIE (N = 44,400)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Category | Variable | Quantiles | Mean | SD |
| Min.  | 25% | 50% | 75% | Max.  |
| Lifetime Unitof Cookie | Observed Age of Cookie on Sampling Day (in days)a | 1 | 1 | 28 | 195 | 467 | 120 | 161 |
| Observed (potentially censored) Lifetime of Cookie (in days) a | 1 | 1 | 94 | 428 | 867 | 228 | 270 |
| Uncensored Lifetime of Cookie (in days) ab | 1 | 1 | 94 | 440 | 1,359 | 298 | 406 |
| Observed Number of Active Days | 1 | 1 | 11 | 90 | 839 | 81 | 144 |
| Share of Observed Active Days per Observed Days | .001 | .143 | .540 | 1.000 | 1.000 | .553 | .393 |
| Observed Number of Ad Impressions c | 1 | 5 | 47 | 971 | 504,999 | 2,369 | 9,977 |
| Observed Number of Ad Impressions per Day | .003 | .896 | 2.172 | 8.373 | 1,017.200 | 9.064 | 24.717 |
| Value of Cookie per LifetimeUnit | Observed Value of Cookie per Day (in €) | .000 | .000 | .001 | .006 | .494 | .005 | .013 |
| Uncensored Value of Cookie per Day (in €) | .000 | .000 | .001 | .006 | .674 | .006 | .016 |
| Observed Value of Cookie per Ad Impression (in €, CPM) | .000 | .389 | .683 | .948 | 54.819 | .720 | .651 |
| Lifetime Valueof Cookie | Observed (potentially censored) Lifetime Value of Cookie (in €) | .000 | .002 | .030 | .717 | 332.710 | 1.569 | 5.459 |
| Predicted Censored Lifetime Value of Cookie for Observed Lifetime (in €) | .000 | .002 | .029 | .799 | 401.728 | 1.791 | 6.280 |
| Mean Absolute Percentage Error (MAPEd) for Observed Lifetime | .000 | .000 | .003 | .102 | 6.278 | .083 | .193 |
| Predicted Residual Lifetime Value of Cookie for Residual Lifetime (in €) | .000 | .000 | .000 | .000 | 167.159 | .960 | 5.242 |
| Uncensored Lifetime Value of Cookie (in €) e | .000 | .002 | .029 | .892 | 453.524 | 2.799 | 11.066 |
| Note: a Rounded to the next full day. b We use a Weibull model to determine the expected residual lifetime for 14,262% of the cookies with potentially censored cookie lifetime. The adjusted average predicted cookie lifetime of 298 days is 30.702% larger than the average observed cookie lifetime of 228 days in the data (i.e., sample 2). c MAPE corresponds to the in-sample absolute difference between the observed lifetime value of the cookie and the predicted lifetime value of the cookie divided by the observed lifetime value of the cookie. e We determine the uncensored cookie lifetime value using the regression outlined in Equation 1 (i.e., model 2 in Table 7).  |

TABLE W7.2
SAMPLES 2: SURVIVAL MODEL PARAMETERS AND FIT MEASURES

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Shape Parameter[95%-CI] | SE | Scale Parameter[95%-CI] | SE | LL | AIC | BIC |
| SAMPLE 2 (N = 44,400) |
| Weibull | .975[.963; .986] | .006 | 466.783[460.434; 473.219] | 3.262 | −154,132.700 | 308,269.500 | 308,285.900 |
| Lognormal | 5.641[5.624; 5.658] | .009 | 1.402[1.388; 1.416] | .007 | −154,630.200 | 309,264.500 | 309,280.900 |
| Notes: To avoid that those cookies with very short lifetimes impact our results too strongly, we only consider cookies with an observed cookie lifetime of seven or more days to predict residual cookie lifetime. CI: confidence interval; SE: standard error; LL: loglikelihood value; AIC: Akaike information criterion; BIC: Bayesian information criterion |

TABLE W7.3
REGRESSION RESULTS OF IMPACT OF TIME ON THE AVERAGE PRICE PER AD IMPRESSION PER DAY

|  |  |  |  |
| --- | --- | --- | --- |
|  | Significant PositiveIncremental Effect | Significant NegativeIncremental Effect | Nonsignificant Incremental Effect |
| Model 1 | Model 2 | Model 1 | Model 2 | Model 1 | Model 2 |
| Dependent Variable(in €; CPM) | Average Price per Ad Impression per Day |
| Constant [95%- CI] | .487[.325; .663] | .609[.402; 845] | 1.130[.923; 1.308] | 1.118[.872; 1.335] | .000[.000; .000] | .000[.000; .000] |
| Time Parameter[95%- CI] | .001[.000; .002] | .001[.000; 002] | −.001[−.002; .000] | −.001[−.002; −.000] | .000[.000; .000] | .000[.000; .000] |
| AdditionalCovariates | N | Y | N | Y | N | Y |
| Number of Cookies(% of all cookies) | 5,890(13.266%) | 6,106(13.752%) | 4,376(9.856%) | 4,097(9.227%) | 34,134(76.878%) | 34,197(77.020%) |
| Total Number of Ad Impressions (% of all) | 49, 095.327(46.669%) | 54,910,981(52.197%) | 19, 926, 405(18.942%) | 17,089,452(16.245%) | 36,177,071(34.389%) | 33,198,370(31.558%) |
| Ad Impressions per Cookie | 8,336 | 8,993 | 4,554 | 4,172 | 1,060 | 971 |
| Mean (Median) Uncensored Cookie Lifetime (in days) | 629(482) | 632(490) | 586(448) | 586(440) | 203(12) | 203(12) |
| Mean (Median) UncensoredCookie Lifetime Value (in €) | 10.123(2.927) | 10.756(3.002) | 4.950(1.081) | 4,761(1.093) | 1.205(.009) | 1.143(.009) |
| Number of Cookies with Significant Zero Effect (% of all cookies) | — | — | — | — | 10,196(22.964%) | 10,132(22.820%) |
| Notes: Unless otherwise noted, this table reports our sample's median estimates from 44,400 cookie-specific regressions. We consider the value increase (decrease) per day to be positive (negative) if the sign of the time parameter is positive (negative) and the value of the parameter is significant (at a 1% level). If the value is insignificant, then we conclude that there is no increase (decrease) in value over time. We apply a small Winsorization to accommodate outliers and replace the most extreme values with the 99% quantile of the respective parameter estimate. Model 1 only includes the time parameter (here: day count) as the independent variable. Model 2 includes the time parameter (here: day count) and additional covariates (i.e., ad inventory characteristics) as independent variables. |

# Web Appendix References

Agencia Española de Protección de Datos (2020), “Guía Sobre el Uso de las Cookies,” Spanish Data Protection Agency (July), <https://www.aepd.es/sites/default/files/2020-07/guia-cookies.pdf>.

Article 29 Data Protection Working Party (2010), “Opinion 2/2010 on Online Behavioural Advertising,” Brussels: European Commission (June 22), <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp171_en.pdf>.

Article 29 Data Protection Working Party (2011), “Opinion 16/2011 on EASA/IAB Best Practice Recommendation on Online Behavioural Advertising,“ Brussels: European Commission (December 8), <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2011/wp188_en.pdf>.

Article 29 Data Protection Working Party (2013), “Working Document 02/2013 providing Guidance on Obtaining Consent for Cookies” (October 2), <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp208_en.pdf>.

Article 29 Data Protection Working Party (2015), “Cookie Sweep Combined Analysis Report“ (February 3), <https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=56123>.

Barker, Alex and Madhumita Murgia (2020), “Google to Phase Out Most Invasive Internet Tracking,” *Financial Times* (January 14), <https://www.ft.com/content/1f56591e-36e1-11ea-a6d3-9a26f8c3cba4>.

Commission Nationale de l'Informatique et des Libertés (2020), “Délibération n° 2020-092 du 17 Septembre 2020 portant Adoption d’une Recommandation Proposant des Modalités Pratiques de mise en Conformité en cas de Recours aux ‘ Cookies et autres Traceurs’,” French Data Protection Authority (September 17), <https://www.cnil.fr/sites/default/files/atoms/files/recommandation-cookies-et-autres-traceurs.pdf>.

Council of the European Union (2021), “Proposal for a Regulation of the European Parliament and of the Council Concerning the Respect for Private Life and the Protection of Personal Data in Electronic Communications and Repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)”, (January 21), https://data.consilium.europa.eu/doc/document/ST-5008-2021-INIT/en/pdf.

Cook, Sam (2017), “How to Comply with Cookie Legislation and Respect Your Website Visitors’ Privacy,” Comparitech (February 24), <https://www.comparitech.com/blog/vpn-privacy/how-to-comply-with-cookie-legislation>.

Cookiebot (2022), “GDPR and Cookies | Compliant Cookie Use with Cookiebot CMP” (January 17), <https://www.cookiebot.com/en/gdpr-cookies/>.

Curac-Dahl, Peter and Marek Juszczyński (2021), “How a Cookie Audit Can Get You Up To Date in Today’s Digital Privacy Landscape,” Piwik Pro (April 15), https://piwik.pro/blog/cookie-audit-digital-privacy.

European Union (2017a), “Proposal for a Regulation of the European Parliament and the Council Concerning the Respect for Private Life and the Protection of Personal Data in Electronic Communications and Repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications),” EUR-Lex (January 10), <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52017PC0010>.

European Union (2017b), “Proposal for a Regulation of the European Parliament and the Council Concerning the Respect for Private Life and the Protection of Personal Data in Electronic Communications and Repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications),” Council of the European Union (September 8), <http://data.consilium.europa.eu/doc/document/ST-11995-2017-INIT/en/pdf>.

Garante per la Protezione dei Dati Personali (2015), “Cookies Instructions Kit,” Italian Data Protection Authority (June 2), <https://help.iubenda.com/wp-content/uploads/2018/04/Cookie-Law-Official-Kit-en.pdf>.

Google (2022), “How Google Retains Data That We Collect” (accessed January 29, 2022), <https://policies.google.com/technologies/retention>.

Koch, Richie (2022), “Cookies, the GDPR, and the ePrivacy Directive” (accessed January 29, 2022), <https://gdpr.eu/cookies/>.

OneTrust (2022), “OneTrust Cookies,” CookiePro (January 10), <https://community.cookiepro.com/s/article/UUID-1e75fb0d-bb79-2af8-cb85-f905d16f1220>.

Sanchez-Rola, Iskander, Matteo Dell’Amico, Platon Kotzias, Leyla Bilge, Pierre-Antoine Vervier, and Igor Santos (2019), “Can I Opt Out Yet? GDPR and the Global Illusion of Cookie Control,” in *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*. New York: Association for Computing Machinery, 340–51.

Tradelab (2019), “Privacy,” (July 19), <https://tradelab.com/en/privacy/>.

Voisin, Gabriel, Ruth, Boardman, Simon Assion, Clara Clark Nevola, Joaquin Munoz Rodriguez (2021), “ICO, CNIL, German and Spanish DPA Revised Cookies Guidelines: Convergence and Divergence,” (May), Bird & Bird, <https://www.twobirds.com/en/news/articles/2019/global/ico-and-cnil-revised-cookies-guidelines-convergence-and-divergence>.

1. With the execption that the lifespan of analytic cookies benefitting from the CNIL consent exemption must not exceed 13 months. [↑](#footnote-ref-2)