Supplementary material for thesis "The use of submodels as a basis for efficient estimation of complex models"

by

Abdollah Safari

M.Sc., University of Tehran, Iran, 2012
B.Sc., University of Tehran, Iran, 2010

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Statistics and Actuarial Science
Faculty of Science

Supplementary material for "Chapter 2 : Parameter-Driven Models for Time Series of Count Data"

Wednesday 20$^{\text{th}}$ September, 2017

1 Models with single covariates

In this section, we summarize the performance of our estimators for different true models in the class (1). We consider the case of a single covariate with different formats including binary, continuous and trend. In each case, we set $\beta_0 = 2$ and $\beta_1 = 0.5$. In addition, we consider two different sample sizes, $n = 100$ and $n = 1000$. For each run, we have 4000 replicates.

Table SM 1: Latent variable distribution : FMM. Binary covariate, $\beta_1 = 0.5$.

n	# of components	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
100	2	medium	low	high	0.060 0.003	0.055 0.005	0.055 0.005
100	3	medium	low	medium	0.060 0.005	0.060 0.003	0.060 0.003
100	4	medium	low	low	0.061 0.003	0.058 0.002	0.057 0.002
1000	2	medium	low	high	0.018 0.000	0.017 0.000	0.017 0.000
1000	3	medium	low	medium	0.020 0.000	0.019 -0.003	0.019 -0.002
1000	4	medium	low	low	0.019 0.0000	0.019 -0.001	0.018 -0.001

Table SM 2: Latent variable distribution : HMM2. Binary covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
100	low	low	medium	0.038 0.001	0.038 0.001	0.038 0.001
100	medium	medium	high	0.057 0.003	0.054 0.005	0.044 0.003
100	high	high	high	0.076 0.004	0.050 0.005	0.042 0.003
1000	low	low	medium	0.012 0.0001	0.012 0.0001	0.012 0.0001
1000	medium	medium	high	0.019 0.001	0.018 0.001	0.014 0.0006
1000	high	high	high	0.026 -0.0002	0.017 0.0005	0.013 0.0003

Table SM 3: Latent variable distribution : HMM3. Binary covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
100	medium	high	medium	0.058 0.0004	0.058 0.0001	0.049 0.0001
100	high	high	high	0.110 -0.001	0.106 0.015	0.081 0.002
1000	medium	high	medium	0.019 0.0001	0.019 -0.002	0.016 -0.001
1000	high	high	high	0.034 -0.001	0.025 0.037	0.018 0.011

Table SM 4: Latent variable distribution : HMM4. Binary covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
100	medium	medium	medium	0.057 -0.001	0.054 -0.001	0.044 0.001
100	high	high	high	0.112 -0.002	0.081 -0.012	0.069 -0.004
1000	medium	medium	medium	0.019 0.001	0.019 -0.001	0.014 -0.0004
1000	high	high	high	0.028 -0.0002	0.020 0.008	0.017 0.003

Table SM 5: Latent variable distribution : HMM10. Binary covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
100	low	medium	medium	0.041 0.003	0.041 0.003	0.040 0.003
100	low	high	medium	0.045 -0.001	0.046 -0.001	0.043 -0.001
100	medium	high	medium	0.051 0.001	0.052 0.001	0.044 0.001
1000	low	medium	medium	0.014 0.001	0.014 0.001	0.013 0.001
1000	low	high	medium	0.015 -0.001	0.015 -0.0003	0.013 -0.0004
1000	medium	high	medium	0.018 -0.001	0.018 -0.002	0.015 -0.0003

Table SM 6: Latent variable distribution : HMM2. Different covariate, $\beta_1 = 0.2$.

n	Covariate	OD	AC1	SP	GLM SSD bias	HMM2 SSD bias
1000	Binary	medium	high	high	0.025 0.000	0.012 -0.000
1000	Continuous	medium	high	medium	0.024 -0.000	0.012 -0.000
1000	Trend	medium	high	medium	0.068 -0.001	0.012 0.000
1000	Monthly seasonal	medium	high	medium	0.029 -0.000	0.012 0.000

Table SM 7: Latent variable distribution : HMM2. Different covariate, $\beta_0 = 2$.

n	Covariate	OD	AC1	SP	GLM SSD bias	HMM2 SSD bias
1000	Binary	medium	high	high	0.069 -0.003	0.087 -0.018
1000	Continuous	medium	high	medium	0.068 -0.004	0.088 -0.021
1000	Trend	medium	high	medium	0.068 -0.006	0.084 -0.016
1000	Monthly seasonal	medium	high	medium	0.068 -0.001	0.087 -0.020

Table SM 8: Latent variable distribution : MA(1). Binary covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
100	medium	medium	low	0.050 0.001	0.051 0.0007	0.050 0.0006
100	high	high	low	0.045 0.0005	0.046 0.0004	0.045 0.0003
1000	medium	medium	low	0.016 -0.0000	0.016 -0.0002	0.015 -0.0001
1000	high	high	low	0.015 0.0000	0.015 0.0000	0.014 -0.0001

Table SM 9: Latent variable distribution : AR(0). Binary covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
1000	low	low	low	0.015 0.001	0.015 0.001	0.015 0.001
1000	medium	low	low	0.016 0.001	0.017 0.001	0.017 0.001
1000	high	low	low	0.024 -0.001	0.029 -0.006	0.030 -0.006

Table SM 10: Latent variable distribution : AR(1). Binary covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
1000	medium	low	low	0.020 -0.002	0.022 -0.004	0.022 -0.003
1000	medium	medium	low	0.019 0.000	0.021 -0.002	0.021 -0.001
1000	medium	high	low	0.019 -0.000	0.021 -0.002	0.018 -0.001
1000	high	high	low	0.026 0.000	0.032 -0.007	0.026 -0.001

Table SM 11: Latent variable distribution : AR(1). Trend covariate, $\beta_1 = 0.5$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
100	medium	high	low	0.536 -0.004	0.541 -0.003	0.632 -0.012
100	high	high	low	0.781 0.015	0.790 0.008	0.876 0.001
1000	medium	high	low	0.183 0.001	0.187 -0.002	0.236 -0.000
1000	high	high	low	0.285 -0.004	0.305 -0.014	0.369 -0.009

Table SM 12: Latent variable distribution : AR(1). Different covariate, $\beta_1 = 0.2$.

n	Covariate	OD	AC1	SP	GLM SSD bias	HMM2 SSD bias
1000	Binary	medium	medium	low	0.025 -0.000	0.030 -0.002
1000	Continuous	medium	medium	low	0.025 0.001	0.030 -0.001
1000	Trend	medium	medium	low	0.040 -0.002	0.054 -0.004
1000	Monthly seasonal	medium	medium	low	0.031 -0.001	0.043 -0.003

Table SM 13: Latent variable distribution : AR(0). Binary covariate, $\beta_0 = 2$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
1000	low	low	low	0.014 1.499	0.015 1.499	0.015 1.499
1000	medium	low	low	0.017 1.500	0.017 1.500	0.017 1.500
1000	high	low	low	0.025 1.500	0.025 1.500	0.026 1.501

Table SM 14: Latent variable distribution : AR(1). Binary covariate, $\beta_0 = 2$.

n	OD	AC1	SP	GLM SSD bias	FMM2 SSD bias	HMM2 SSD bias
1000	medium	low	low	0.023 0.000	0.024 0.001	0.024 0.001
1000	medium	medium	low	0.028 -0.001	0.028 -0.000	0.028 -0.000
1000	medium	high	low	0.053 0.004	0.053 -0.004	0.053 0.005
1000	high	high	low	0.085 -0.005	0.085 -0.005	0.089 -0.001

Table SM 15: Latent variable distribution : AR(1). Different covariate, $\beta_0 = 2$.

n	Covariate	OD	AC1	SP	GLM SSD bias	HMM2 SSD bias
1000	Binary	medium	medium	low	0.038 -0.002	0.041 -0.000
1000	Continuous	medium	medium	low	0.039 -0.003	0.039 -0.002
1000	Trend	medium	medium	low	0.038 -0.003	0.041 -0.000
1000	Monthly seasonal	medium	medium	low	0.040 -0.002	0.042 -0.001

Table SM 16: Most efficient estimators for different levels of the factors when the model has one covariate.

	Factor			Estimator	
Run	OD	AC1	SP	GLM	HMM2
1	Low	Low	Low	*	
2	Low	Low	Medium	*	
3	Low	Medium	Low [+]	-	-
4	Low	Medium	Medium		*
5	Medium	Low	Low	*	
6	Medium	Medium	Low	*	
7	Medium	Low	Medium [+]	-	-
8	Medium	High	Low	*	
9	Medium	Medium	Medium		*
10	Medium	Low	High		*
11	Medium	High	Medium		*
12	Medium	Medium	High		*
13	Medium	High	High [+]	-	-
14	High	Low	Low	*	
15	High	Low	Medium	*	
16	High	Medium	Low	*	
17	High	High	Low	*	
18	High	Medium	Medium		*
19	High	Low	High [+]	-	-
20	High	High	Medium		*
21	High	Medium	High [+]	-	-
22	High	High	High		*

[+] Not a possible combination

8

Table SM 17: Latent variable distribution : HMM2. Covariate: Trend and seasonal. $\beta_0 = 2$ and $\beta_1 = \beta_2 = 0.1$.

OD	AC1	SP	Coefficient	GLM SSD bias	HMM2 SSD bias
medium	high	medium	β_0	0.054	0.080
				0.001	-0.019
			β_1 (trend)	0.055	0.013
				-0.000	-0.000
			β_2 (seasonal)	0.025	0.014
				0.000	-0.000

Table SM 18: Latent variable distribution : HMM2. Covariate: Trend and binary. $\beta_0 = 2$ and $\beta_1 = \beta_2 = 0.1$.

OD	AC1	SP	Coefficient	GLM SSD bias	HMM2 SSD bias
medium	high	medium	β_0	0.056	0.079
				-0.002	-0.026
			β_1 (trend)	0.056	0.013
				0.000	-0.000
			β_2 (binary)	0.021	0.013
				-0.000	0.000

2   Models with multiple covariates

In this section, we consider the case of multiple covariates with different formats including binary, continuous and trend. In each case, we set $\beta_0 = 2$ and $\beta_i = 0.1$ (for all the covariates). In addition, we only consider the sample size $n = 1000$. For each run, we have 4000 replicates. Since we never suggest the FMM estimator, we present only the results of the GLM and HMM estimators in this section.

Table SM 19: Latent variable distribution : HMM2. Covariate: Binary and binary. $\beta_0 = 2$ and $\beta_1 = \beta_2 = 0.1$.

OD	AC1	SP	Coefficient	GLM SSD bias	HMM2 SSD bias
medium	high	high	β_0	0.058	0.077
				-0.004	-0.030
			β_1 (binary)	0.023	0.013
				0.001	0.001
			β_2 (binary)	0.020	0.013
				0.000	-0.000

Table SM 20: Latent variable distribution : HMM2. Covariate: Trend, continuous, seasonal, and binary. $\beta_0 = 2$ and $\beta_i = 0.1$, $i = 1, \ldots, 4$.

OD	AC1	SP	Coefficient	GLM SSD bias	HMM2 SSD bias
medium	high	medium	β_0	0.069	0.144
				-0.007	-0.119
			β_1 (trend)	0.068	0.015
				0.001	-0.000
			β_2 (continuous)	0.024	0.016
				-0.000	-0.000
			β_3 (seasonal)	0.029	0.016
				0.000	-0.000
			β_4 (binary)	0.026	0.016
				-0.001	-0.000

Table SM 21: Latent variable distribution : HMM2. Different covariate, $\beta_1 = 0.2$.

Covariate	OD	AC1	SP	GLM SSD White DDW *	HMM2 SSD White
Binary	medium	high	high	0.025 0.026 0.025	0.012 0.012
Continuous	medium	high	medium	0.024 0.025 0.024	0.012 0.012
Trend	medium	high	medium	0.068 0.039 0.034	0.012 0.012
Monthly seasonal	medium	high	medium	0.029 0.038 0.017	0.012 0.012

* Davis, Dunsmuir and Wang (2000)

3   Standard error estimate of the estimators

In this section, we present the results of the standard error (SE) estimate of the GLM and HMM estimators. We consider variety of cases including single and multiple covariates, both with different covariate formats. Since we estimate the asymptotic SE of the estimators, we only consider the sample size $n = 1000$. For each run, we have 4000 replicates. We treat the sample standard deviation of the estimators as the true SE and then compare the SSDs to the average SEs of the estimators.

Table SM 22: Latent variable distribution : AR(1). Different covariate, $\beta_1 = 0.2$.

Covariate	OD	AC1	SP	GLM SSD White DDW	HMM2 SSD White
Binary	medium	medium	low	0.025 0.025 0.024	0.030 0.030
Continuous	medium	medium	low	0.025 0.025 0.024	0.030 0.030
Trend	medium	medium	low	0.040 0.033 0.019	0.054 0.046
Monthly seasonal	medium	medium	low	0.031 0.032 0.017	0.043 0.043

Table SM 23: Latent variable distribution : HMM2. Different covariate, $\beta_0 = 2$.

Covariate	OD	AC1	SP	GLM SSD White DDW	HMM2 SSD White
Binary	medium	high	high	0.069 0.039 0.004	0.087 0.069
Continuous	medium	high	medium	0.068 0.039 0.004	0.088 0.070
Trend	medium	high	medium	0.068 0.039 NA*	0.084 0.039
Monthly seasonal	medium	high	medium	0.068 0.039 0.003	0.087 0.070

* Convergence issues

Table SM 24: Latent variable distribution : AR(1). Different covariate, $\beta_0 = 2$.

Covariate	OD	AC1	SP	GLM SSD White DDW	HMM2 SSD White
Binary	medium	medium	low	0.038 0.033 0.002	0.041 0.035
Continuous	medium	medium	low	0.039 0.033 0.002	0.039 0.035
Trend	medium	medium	low	0.038 0.033 0.001	0.041 0.034
Monthly seasonal	medium	medium	low	0.040 0.033 0.003	0.042 0.035

Table SM 25: Latent variable distribution : HMM2. Covariate: Trend and seasonal. $\beta_0 = 2$ and $\beta_1 = \beta_2 = 0.1$.

OD	AC1	SP	Coefficient	GLM SSD White DDW	HMM2 SSD White
medium	high	medium	β_0	0.054 0.032 0.001	0.080 0.062
			β_1 (trend)	0.055 0.032 0.027	0.013 0.013
			β_2 (seasonal)	0.025 0.030 0.015	0.014 0.013

Table SM 26: Latent variable distribution : HMM2. Covariate: Trend and binary. $\beta_0 = 2$ and $\beta_1 = \beta_2 = 0.1$.

OD	AC1	SP	Coefficient	GLM SSD White DDW	HMM2 SSD White
medium	high	medium	β_0	0.056 0.032 0.002	0.079 0.064
			β_1 (trend)	0.056 0.032 0.027	0.013 0.013
			β_2 (binary)	0.021 0.022 0.021	0.013 0.012

Table SM 27: Latent variable distribution : HMM2. Covariate: Binary and binary. $\beta_0 = 2$ and $\beta_1 = \beta_2 = 0.1$.

OD	AC1	SP	Coefficient	GLM SSD White DDW	HMM2 SSD White
medium	high	high	β_0	0.058 0.032 0.003	0.077 0.060
			β_1 (trend)	0.023 0.023 0.022	0.013 0.012
			β_2 (binary)	0.020 0.021 0.020	0.013 0.012

14

Table SM 28: Latent variable distribution : HMM2. Covariate: Trend, continuous, seasonal, and binary. $\beta_0 = 2$ and $\beta_i = 0.1$, $i = 1, \ldots, 4$.

OD	AC1	SP	Coefficient	GLM SSD White DDW	HMM2 SSD White
			β_0	0.069 0.038 0.003	0.144 0.127
			β_1 (trend)	0.068 0.038 0.033	0.015 0.012
medium	high	medium	β_2 (continuous)	0.024 0.024 0.024	0.016 0.012
			β_3 (seasonal)	0.029 0.036 0.017	0.016 0.012
			β_4 (binary)	0.026 0.025 0.025	0.016 0.012

Supplementary material for "Chapter 3 : Display advertising: Estimating conversion probability efficiently"

Wednesday 20th September, 2017

1 Delay distribution given covariates

We use the campaign with $ID = 95667a0f$. We fitted an exponential and a Weibull model on the transformed (Box-Cox transformation) delays of converted clicks of the chosen campaign. Figure SM**??** shows the estimated survival function for each model, where we can the Weibull estimate is closer to that of the KM.
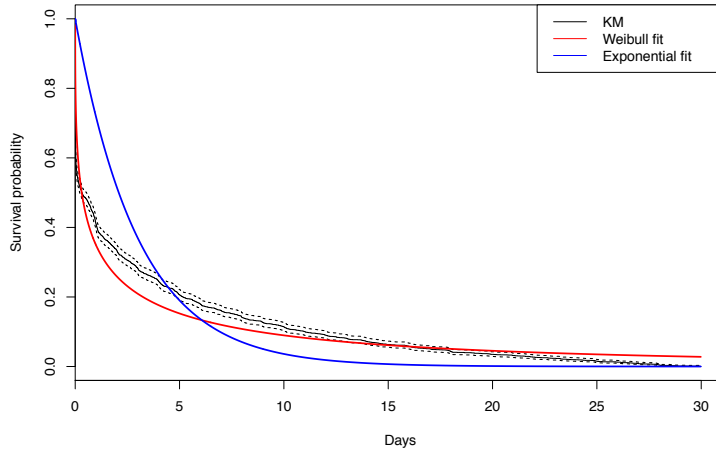
Fig. SM 1: Estimated survival probability of Weibull and Exponential models comparing with the KM probabilities

2 True covariate coefficients in the delay and conversion models

We fit a logistic regression model to the eventual conversion status of the clicks from the chosen campaign, and also a Weibull model to the delays of the converted clicks of the same campaign. We then use the estimated coefficients as the true covariate coefficients in our simulation study. Table SM**??** shows the estimated covariate coefficient values from each model:

In study 2, we used a fixed shape parameter equal to 0.5 in all runs.

3 Average bias of the estimators

3.1 Study 1

Table SM**??** shows the average bias of the estimators in study 1 where we generate exponential distributed conversion delays.

3.2 Study 2

Table SM**??** shows the average bias of the estimators in study 2 where we generate Weibull distributed conversion delays.

Table SM 1: True covariate coefficient values used in the simulation study.

Covariates	Coefficients	
	Delay model	Conversion probability model
V4	0.28	0.02
V8	-0.004	-0.27
V9	-0.01	0.11
V10	-0.33	0.19
V11 1c9284e8	0.60	-0.33
V11 3961c76e	0.44	-0.20
V11 46c13849	-0.59	-0.21
V11 47539da2	-0.03	0.01
V11 8cdde828	0.22	-0.003
V11 917117e2	-0.13	-0.04
V11 c76c0623	0.43	-0.17
V11 d208d2cc	-0.05	-0.09
V14 5d8e5fa7	0.41	0.10
V14 5f111c42	-0.06	0.90
V14 917de545	0.24	0.98
V14 c2fbb0b7	-0.10	-0.90
V14 e35e0e8f	0.40	0.02
V18 3486227d	-0.09	-0.09
V18 8efede7f	0.17	0.06

Table SM 2: Average bias of the estimators over time when the delays follow an exponential distribution.

Delay time	Conversion probability	Estimators	Days since the first click						
			10	15	21	30	39	51	60
Low	Low	True-E-adjusted	-0.01	-0.006	-0.005	-0.004	-0.003	-0.003	-0.002
		E-Bias-adjusted	-0.004	0.011	0.017	0.024	0.017	0.01	0.009
		DFM	0.021	-0.069	-0.081	-0.072	-0.086	-0.08	-0.086
		Naive	-0.022	-0.014	-0.012	-0.008	-0.006	-0.006	-0.004
Low	Mod	True-E-adjusted	-0.028	-0.018	-0.016	-0.011	-0.008	-0.006	-0.004
		E-Bias-adjusted	-0.025	0.008	0.023	0.042	0.029	0.016	0.012
		DFM	-0.016	-0.016	-0.020	-0.017	-0.018	-0.017	-0.019
		Naive	-0.064	-0.043	-0.035	-0.024	-0.017	-0.014	-0.012
Low	High	True-E-adjusted	-0.044	-0.028	-0.027	-0.016	-0.009	-0.01	-0.009
		E-Bias-adjusted	-0.07	-0.019	0.001	0.028	0.021	0.007	0.003
		DFM	0.062	0.045	0.061	0.013	0.012	0.014	0.011
		Naive	-0.118	-0.077	-0.063	-0.042	-0.028	-0.025	-0.022
Mod	Low	True-E-adjusted	-0.013	-0.009	-0.008	-0.005	-0.004	-0.004	-0.003
		E-Bias-adjusted	-0.024	-0.006	0.003	0.014	0.01	0.005	0.004
		DFM	0.241	-0.059	-0.07	-0.071	-0.087	-0.084	-0.088
		Naive	-0.039	-0.028	-0.023	-0.016	-0.013	-0.01	-0.009
Mod	High	True-E-adjusted	-0.053	-0.037	-0.035	-0.023	-0.016	-0.016	-0.015
		E-Bias-adjusted	-0.158	-0.084	-0.049	-0.007	-0.005	-0.012	-0.014
		DFM	-0.116	-0.103	-0.069	0.063	0.057	0.098	0.069
		Naive	-0.207	-0.147	-0.118	-0.083	-0.06	-0.05	-0.045
High	Low	True-E-adjusted	-0.012	-0.008	-0.008	-0.005	-0.004	-0.004	-0.004
		E-Bias-adjusted	-0.043	-0.023	-0.012	0.003	0.001	-0.001	-0.002
		DFM	0.045	-0.029	-0.073	-0.065	-0.083	-0.084	-0.089
		Naive	-0.055	-0.042	-0.035	-0.025	-0.02	-0.016	-0.015
High	Mod	True-E-adjusted	-0.033	-0.021	-0.019	-0.013	-0.011	-0.011	-0.01
		E-Bias-adjusted	-0.127	-0.075	-0.045	-0.009	-0.009	-0.013	-0.014
		DFM	0.051	-0.227	-0.243	-0.199	-0.203	-0.197	-0.221
		Naive	-0.156	-0.119	-0.097	-0.071	-0.056	-0.046	-0.042
High	High	True-E-adjusted	-0.068	-0.048	-0.044	-0.029	-0.023	-0.025	-0.025
		E-Bias-adjusted	-0.255	-0.166	-0.116	-0.057	-0.044	-0.043	-0.043
		DFM	-0.235	-0.288	-0.234	-0.061	-0.043	0.026	-0.009
		Naive	-0.3	-0.228	-0.187	-0.137	-0.106	-0.088	-0.081

Table SM 3: Average bias of the estimators over time when the delays follow a Weibull distribution.

Delay time	Conversion probability	Estimators	Days since the first click						
			10	15	21	30	39	51	60
Low	Low	E-Bias-adjusted	-0.011	0.001	0.009	0.017	0.011	0.006	0.004
		DFM	0.003	-0.075	-0.084	-0.077	-0.086	-0.081	-0.087
		Naive	-0.028	-0.022	-0.018	-0.014	-0.011	-0.009	-0.008
Low	Mod	E-Bias-adjusted	-0.041	-0.011	0.006	0.027	0.015	0.006	0.002
		DFM	-0.165	-0.203	-0.221	-0.188	-0.196	-0.183	-0.200
		Naive	-0.078	-0.061	-0.05	-0.038	-0.031	-0.025	-0.023
Low	High	E-Bias-adjusted	-0.102	-0.053	-0.025	0.006	0.001	-0.01	-0.014
		DFM	-0.02	-0.025	0.002	0.083	0.075	0.098	0.07
		Naive	-0.148	-0.11	-0.09	-0.065	-0.05	-0.045	-0.043
Mod	Low	E-Bias-adjusted	-0.025	-0.012	-0.004	0.006	0.003	-0.002	-0.003
		DFM	0.048	-0.062	-0.083	-0.08	-0.088	-0.083	-0.088
		Naive	-0.04	-0.033	-0.028	-0.023	-0.019	-0.017	-0.015
Mod	High	E-Bias-adjusted	-0.168	-0.111	-0.079	-0.04	-0.04	-0.046	-0.048
		DFM	-0.15	-0.163	-0.157	-0.027	-0.022	0.027	-0.011
		Naive	-0.215	-0.172	-0.148	-0.117	-0.096	-0.085	-0.081
High	Low	E-Bias-adjusted	-0.036	-0.023	-0.015	-0.003	-0.006	-0.009	-0.01
		DFM	0.028	-0.064	-0.081	-0.083	-0.086	-0.089	-0.091
		Naive	-0.049	-0.042	-0.037	-0.03	-0.026	-0.023	-0.022
High	Mod	E-Bias-adjusted	-0.111	-0.078	-0.056	-0.027	-0.031	-0.036	-0.037
		DFM	-0.207	-0.241	-0.258	-0.229	-0.218	-0.224	-0.233
		Naive	-0.142	-0.12	-0.106	-0.087	-0.075	-0.066	-0.063
High	High	E-Bias-adjusted	-0.221	-0.162	-0.126	-0.082	-0.079	-0.083	-0.085
		DFM	-0.258	-0.285	-0.271	-0.134	-0.134	-0.083	-0.116
		Naive	-0.267	-0.222	-0.195	-0.16	-0.137	-0.124	-0.12

Table SM 4: SE of the estimators over time when the delays follow an exponential distribution.

Delay time	Conversion probability	Estimators	Days since the first click						
			10	15	21	30	39	51	60
Low	Low	True-E-adjusted	0.015	0.014	0.013	0.013	0.012	0.012	0.004
		E-Bias-adjusted	0.020	0.019	0.019	0.019	0.017	0.015	0.006
		Naive	0.012	0.012	0.012	0.012	0.012	0.011	0.004
Low	Mod	True-E-adjusted	0.029	0.028	0.027	0.027	0.025	0.015	0.006
		E-Bias-adjusted	0.039	0.040	0.040	0.041	0.036	0.021	0.009
		Naive	0.024	0.024	0.024	0.025	0.024	0.014	0.006
Low	High	True-E-adjusted	0.017	0.016	0.015	0.014	0.014	0.009	0.005
		E-Bias-adjusted	0.014	0.012	0.010	0.008	0.007	0.006	0.006
		Naive	0.017	0.017	0.017	0.015	0.015	0.09	0.005
Mod	Low	True-E-adjusted	0.016	0.014	0.013	0.013	0.012	0.012	0.005
		E-Bias-adjusted	0.017	0.017	0.017	0.017	0.016	0.015	0.007
		Naive	0.011	0.011	0.011	0.011	0.011	0.011	0.005
Mod	High	True-E-adjusted	0.018	0.016	0.015	0.013	0.013	0.009	0.005
		E-Bias-adjusted	0.016	0.014	0.011	0.008	0.007	0.007	0.006
		Naive	0.015	0.016	0.016	0.015	0.015	0.009	0.005
High	Low	True-E-adjusted	0.017	0.015	0.0137	0.013	0.012	0.012	0.007
		E-Bias-adjusted	0.014	0.015	0.015	0.0156	0.015	0.014	0.007
		Naive	0.009	0.010	0.010	0.010	0.010	0.010	0.006
High	Mod	True-E-adjusted	0.031	0.029	0.028	0.027	0.026	0.0176	0.009
		E-Bias-adjusted	0.024	0.028	0.031	0.033	0.031	0.021	0.011
		Naive	0.016	0.018	0.019	0.021	0.021	0.015	0.007
High	High	True-E-adjusted	0.021	0.017	0.014	0.0123	0.0123	0.009	0.006
		E-Bias-adjusted	0.017	0.016	0.012	0.009	0.009	0.008	0.007
		Naive	0.013	0.014	0.014	0.014	0.014	0.010	0.005

4 SE of the estimators

4.1 Study 1

Table SM?? shows the SE of the estimators in study 1 where we generate exponential distributed conversion delays.

4.2 Study 2

Table SM?? shows the SE of the estimators in study 2 where we generate Weibull distributed conversion delays.

Table SM 5: SE of the estimators over time when the delays follow a Weibull distribution.

Delay time	Conversion probability	Estimators	Days since the first click						
			10	15	21	30	39	51	60
Low	Low	E-Bias-adjusted	0.017	0.017	0.017	0.017	0.016	0.014	0.006
		Naive	0.011	0.011	0.011	0.011	0.011	0.010	0.004
Low	Mod	E-Bias-adjusted	0.037	0.037	0.038	0.039	0.035	0.021	0.009
		Naive	0.023	0.023	0.023	0.024	0.023	0.014	0.005
Low	High	E-Bias-adjusted	0.017	0.013	0.011	0.008	0.009	0.010	0.008
		Naive	0.017	0.013	0.011	0.008	0.009	0.010	0.008
Mod	Low	E-Bias-adjusted	0.016	0.016	0.016	0.016	0.015	0.014	0.006
		Naive	0.016	0.016	0.016	0.016	0.015	0.014	0.006
Mod	High	E-Bias-adjusted	0.019	0.014	0.012	0.009	0.008	0.008	0.007
		Naive	0.019	0.014	0.012	0.009	0.008	0.008	0.007
High	Low	E-Bias-adjusted	0.0154	0.015	0.015	0.015	0.014	0.013	0.006
		Naive	0.010	0.009	0.009	0.010	0.010	0.010	0.005
High	Mod	E-Bias-adjusted	0.029	0.029	0.030	0.031	0.029	0.019	0.009
		Naive	0.018	0.018	0.019	0.019	0.020	0.013	0.006
High	High	E-Bias-adjusted	0.025	0.020	0.018	0.014	0.014	0.013	0.008
		Naive	0.020	0.020	0.022	0.022	0.022	0.014	0.006

Supplementary material for "Chapter 4 : Conversion probability estimation in display advertising: Incorporating information from multiple visits to the same ad"

Wednesday 20$^{\text{th}}$ September, 2017

1 True covariate coefficients in the delay and conversion models

We use the campaign with $ID = 95667a0f$. We use the estimated coefficients of the models from Safari et al. (2017) as the true covariate coefficients in our simulation study. Table SM1 shows the estimated covariate coefficient values from each model:

In study 2, we used a fixed shape parameter equal to 0.5 in all runs.

2 Average bias of the estimators

2.1 Study 1

Tables SM2–SM4 show the average bias of the estimators in study 1 where we generate exponential distributed the S_i's and inter-visit.

2.2 Study 2

Tables SM5–SM7 show the average bias of the estimators in study 2 where we generate Weibull distributed the S_i's and inter-visit times.

3 SE of the estimators

3.1 Study 1

Tables SM8–SM10 show the SE of the estimators in study 1 where we generate exponential distributed the S_i's and inter-visit times.

Table SM 1: True covariate coefficient values used in the simulation study.

| Covariates | Coefficients | |
| | Delay model | Conversion probability model |
		and number of visits model
V4	0.28	0.02
V8	-0.004	-0.27
V9	-0.01	0.11
V10	-0.33	0.19
V11 1c9284e8	0.60	-0.33
V11 3961c76e	0.44	-0.20
V11 46c13849	-0.59	-0.21
V11 47539da2	-0.03	0.01
V11 8cdde828	0.22	-0.003
V11 917117e2	-0.13	-0.04
V11 c76c0623	0.43	-0.17
V11 d208d2cc	-0.05	-0.09
V14 5d8e5fa7	0.41	0.10
V14 5f111c42	-0.06	0.90
V14 917de545	0.24	0.98
V14 c2fbb0b7	-0.10	-0.90
V14 e35e0e8f	0.40	0.02
V18 3486227d	-0.09	-0.09
V18 8efede7f	0.17	0.06

Table SM 2: Average bias of the estimators over time when the S_i's and inter-visit times follow an exponential distribution times – Low probability of conversion.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Low	Low	Low	True-DV-adjusted	0.003	0.025	0.046	0.069	0.071	0.081	0.091
			DV-adjusted	0.022	0.028	0.041	0.051	0.099	0.083	0.066
			E-adjusted	-0.04	-0.022	-0.005	0.015	0.016	0.024	0.031
			Naive	-0.094	-0.07	-0.056	-0.038	-0.026	-0.023	-0.02
Low	Low	Mod	True-DV-adjusted	-0.17	-0.158	-0.135	-0.093	-0.041	-0.012	-0.003
			DV-adjusted	-0.159	-0.16	-0.147	-0.117	-0.072	-0.045	-0.038
			E-adjusted	-0.197	-0.181	-0.159	-0.124	-0.08	-0.056	-0.049
			Naive	-0.215	-0.196	-0.173	-0.139	-0.097	-0.08	-0.072
Low	Low	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.204	-0.234	-0.257	-0.252	-0.176	-0.14	-0.119
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13
Low	Mod	Low	True-DV-adjusted	0.091	0.084	0.072	0.069	0.046	0.025	0.003
			DV-adjusted	0.083	0.067	0.066	0.051	0.041	0.028	0.022
			E-adjusted	0.091	0.084	0.072	0.069	0.046	0.025	0.021
			Naive	-0.095	-0.07	-0.056	-0.038	-0.027	-0.023	-0.021
Low	Mod	Mod	True-DV-adjusted	-0.17	-0.158	-0.136	-0.094	-0.042	-0.012	-0.003
			DV-adjusted	-0.159	-0.16	-0.148	-0.117	-0.072	-0.045	-0.038
			E-adjusted	-0.197	-0.181	-0.16	-0.124	-0.08	-0.056	-0.049
			Naive	-0.215	-0.196	-0.173	-0.139	-0.097	-0.08	-0.072
Low	Mod	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.204	-0.234	-0.257	-0.252	-0.176	-0.14	-0.119
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13
Low	High	Low	True-DV-adjusted	0.091	0.084	0.073	0.07	0.047	0.025	0.003
			DV-adjusted	0.066	0.083	0.102	0.051	0.042	0.028	0.022
			E-adjusted	-0.041	-0.022	-0.004	0.015	0.018	0.026	0.031
			Naive	-0.095	-0.07	-0.056	-0.038	-0.026	-0.023	-0.02
Low	High	Mod	True-DV-adjusted	-0.17	-0.158	-0.135	-0.093	-0.042	-0.012	-0.003
			DV-adjusted	-0.159	-0.16	-0.147	-0.117	-0.072	-0.045	-0.038
			E-adjusted	-0.197	-0.181	-0.159	-0.124	-0.08	-0.056	-0.049
			Naive	-0.215	-0.196	-0.173	-0.139	-0.097	-0.08	-0.072
Low	High	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.204	-0.234	-0.257	-0.252	-0.176	-0.14	-0.119
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13

Table SM 3: Average bias of the estimators over time when the S_i's and inter-visit times follow an exponential distribution times – Moderate probability of conversion.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Mod	Low	Low	True-DV-adjusted	0.091	0.084	0.073	0.07	0.047	0.025	0.003
			DV-adjusted	0.066	0.083	0.101	0.051	0.042	0.028	0.022
			E-adjusted	-0.055	-0.037	-0.02	-0.003	-0.001	0.007	0.012
			Naive	-0.095	-0.07	-0.056	-0.038	-0.026	-0.023	-0.02
Mod	Low	Mod	True-DV-adjusted	-0.17	-0.158	-0.136	-0.094	-0.042	-0.012	-0.003
			DV-adjusted	-0.159	-0.16	-0.147	-0.117	-0.072	-0.045	-0.038
			E-adjusted	-0.205	-0.188	-0.166	-0.132	-0.092	-0.07	-0.063
			Naive	-0.215	-0.196	-0.173	-0.139	-0.097	-0.08	-0.072
Mod	Low	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.217	-0.243	-0.261	-0.255	-0.18	-0.146	-0.126
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13
Mod	Mod	Low	True-DV-adjusted	0.091	0.081	0.071	0.069	0.046	0.025	0.003
			DV-adjusted	0.066	0.083	0.099	0.051	0.041	0.028	0.022
			E-adjusted	-0.055	-0.037	-0.021	-0.003	-0.002	0.005	0.012
			Naive	-0.094	-0.07	-0.056	-0.038	-0.026	-0.023	-0.02
Mod	Mod	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.217	-0.243	-0.261	-0.255	-0.18	-0.146	-0.126
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13
Mod	High	Low	True-DV-adjusted	0.091	0.082	0.071	0.07	0.046	0.026	0.004
			DV-adjusted	0.067	0.083	0.099	0.051	0.041	0.029	0.022
			E-adjusted	-0.054	-0.036	-0.02	-0.003	-0.002	0.005	0.012
			Naive	-0.094	-0.069	-0.056	-0.038	-0.026	-0.023	-0.02
Mod	High	Mod	True-DV-adjusted	-0.169	-0.154	-0.133	-0.092	-0.041	-0.011	-0.003
			DV-adjusted	-0.158	-0.156	-0.145	-0.116	-0.072	-0.044	-0.037
			E-adjusted	-0.205	-0.184	-0.164	-0.131	-0.091	-0.068	-0.062
			Naive	-0.215	-0.192	-0.171	-0.138	-0.096	-0.08	-0.072
Mod	High	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.217	-0.243	-0.261	-0.255	-0.18	-0.146	-0.126
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13

Table SM 4: Average bias of the estimators over time when the S_i's and inter-visit times follow an exponential distribution times – High probability of conversion.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
High	Low	Low	True-DV-adjusted	0.093	0.086	0.077	0.072	0.05	0.029	0.009
			DV-adjusted	0.068	0.085	0.106	0.053	0.044	0.031	0.028
			E-adjusted	-0.05	-0.034	-0.018	-0.001	0.002	0.008	0.014
			Naive	-0.09	-0.068	-0.054	-0.036	-0.025	-0.022	-0.019
High	Low	Mod	True-DV-adjusted	-0.168	-0.154	-0.133	-0.093	-0.042	-0.012	-0.003
			DV-adjusted	-0.158	-0.156	-0.145	-0.116	-0.072	-0.045	-0.037
			E-adjusted	-0.205	-0.184	-0.164	-0.131	-0.092	-0.069	-0.062
			Naive	-0.214	-0.193	-0.171	-0.138	-0.096	-0.08	-0.072
High	Low	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.217	-0.243	-0.261	-0.255	-0.18	-0.146	-0.126
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13
High	Mod	Low	True-DV-adjusted	0.093	0.086	0.077	0.072	0.05	0.029	0.009
			DV-adjusted	0.068	0.085	0.106	0.053	0.044	0.031	0.028
			E-adjusted	-0.05	-0.034	-0.018	-0.001	0.002	0.008	0.014
			Naive	-0.09	-0.068	-0.054	-0.036	-0.025	-0.022	-0.019
High	Mod	Mod	True-DV-adjusted	-0.169	-0.155	-0.133	-0.092	-0.042	-0.012	-0.003
			DV-adjusted	-0.158	-0.157	-0.145	-0.116	-0.072	-0.045	-0.037
			E-adjusted	-0.205	-0.185	-0.164	-0.131	-0.092	-0.069	-0.062
			Naive	-0.215	-0.193	-0.171	-0.138	-0.096	-0.08	-0.072
High	Mod	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.217	-0.243	-0.261	-0.255	-0.18	-0.146	-0.126
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13
High	High	Low	True-DV-adjusted	0.093	0.086	0.077	0.072	0.05	0.029	0.009
			DV-adjusted	0.068	0.085	0.106	0.053	0.044	0.031	0.028
			E-adjusted	-0.05	-0.034	-0.018	-0.001	0.002	0.008	0.014
			Naive	-0.09	-0.068	-0.054	-0.036	-0.025	-0.022	-0.019
High	High	Mod	True-DV-adjusted	-0.168	-0.154	-0.132	-0.092	-0.041	-0.012	-0.003
			DV-adjusted	-0.158	-0.156	-0.144	-0.115	-0.072	-0.045	-0.037
			E-adjusted	-0.205	-0.184	-0.163	-0.131	-0.091	-0.069	-0.062
			Naive	-0.214	-0.193	-0.17	-0.138	-0.096	-0.08	-0.072
High	High	High	True-DV-adjusted	-0.138	-0.197	-0.239	-0.241	-0.156	-0.114	-0.089
			DV-adjusted	-0.097	-0.173	-0.228	-0.239	-0.169	-0.136	-0.117
			E-adjusted	-0.217	-0.243	-0.261	-0.255	-0.18	-0.146	-0.126
			Naive	-0.225	-0.249	-0.265	-0.258	-0.184	-0.149	-0.13

Table SM 5: Average bias of the estimators over time when the S_i's and inter-visit times follow a Weibull distribution – Low conversion probability

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Low	Low	Low	DV-adjusted	0.064	0.067	0.051	0.013	-0.012	-0.004	0.005
			E-adjusted	0.023	0.02	0.012	0.003	-0.045	-0.03	-0.017
			Naive	-0.1	-0.079	-0.068	-0.049	-0.034	-0.029	-0.026
Low	Low	Mod	DV-adjusted	-0.125	-0.131	-0.141	-0.114	-0.069	-0.049	-0.037
			E-adjusted	-0.171	-0.16	-0.156	-0.123	-0.078	-0.061	-0.049
			Naive	-0.193	-0.179	-0.172	-0.14	-0.097	-0.08	-0.072
Low	Low	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124
Low	Mod	Low	DV-adjusted	0.08	0.095	0.089	0.039	0.029	0.02	0.015
			E-adjusted	-0.045	-0.03	-0.016	0.003	0.01	0.019	0.024
			Naive	-0.1	-0.078	-0.067	-0.049	-0.034	-0.029	-0.026
Low	Mod	Mod	DV-adjusted	-0.125	-0.131	-0.141	-0.114	-0.069	-0.049	-0.037
			E-adjusted	-0.171	-0.16	-0.156	-0.123	-0.078	-0.06	-0.049
			Naive	-0.193	-0.179	-0.171	-0.139	-0.097	-0.08	-0.072
Low	Mod	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.13	-0.2	-0.229	-0.226	-0.144	-0.103	-0.081
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124
Low	High	Low	DV-adjusted	0.063	0.077	0.073	0.024	0.015	0.006	0.001
			E-adjusted	-0.059	-0.044	-0.03	-0.012	-0.005	0.003	0.008
			Naive	-0.1	-0.078	-0.067	-0.048	-0.034	-0.028	-0.026
Low	High	Mod	DV-adjusted	-0.083	-0.099	-0.114	-0.083	-0.03	-0.006	0.008
			E-adjusted	-0.171	-0.16	-0.156	-0.123	-0.078	-0.06	-0.049
			Naive	-0.193	-0.179	-0.171	-0.139	-0.097	-0.08	-0.072
Low	High	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124

Table SM 6: Average bias of the estimators over time when the S_i's and inter-visit times follow a Weibull distribution – Moderate conversion probability

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Mod	Low	Low	DV-adjusted	0.08	0.095	0.089	0.039	0.029	0.02	0.015
			E-adjusted	-0.045	-0.031	-0.016	0.003	0.01	0.019	0.024
			Naive	-0.1	-0.078	-0.067	-0.048	-0.034	-0.028	-0.026
Mod	Low	Mod	DV-adjusted	-0.083	-0.099	-0.114	-0.083	-0.03	-0.006	0.009
			E-adjusted	-0.171	-0.16	-0.156	-0.123	-0.078	-0.06	-0.049
			Naive	-0.193	-0.179	-0.171	-0.139	-0.097	-0.08	-0.072
Mod	Low	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124
Mod	Mod	Low	DV-adjusted	0.079	0.096	0.092	0.039	0.029	0.02	0.015
			E-adjusted	-0.045	-0.03	-0.017	0.003	0.012	0.02	0.023
			Naive	-0.099	-0.079	-0.068	-0.049	-0.034	-0.029	-0.026
Mod	Mod	Mod	DV-adjusted	-0.088	-0.103	-0.114	-0.086	-0.035	-0.004	0.007
			E-adjusted	-0.174	-0.162	-0.156	-0.125	-0.082	-0.058	-0.05
			Naive	-0.196	-0.181	-0.171	-0.141	-0.098	-0.081	-0.073
Mod	Mod	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124
Mod	High	Low	DV-adjusted	0.079	0.095	0.093	0.038	0.029	0.019	0.017
			E-adjusted	-0.043	-0.031	-0.017	0.003	0.013	0.02	0.023
			Naive	-0.098	-0.079	-0.068	-0.049	-0.034	-0.029	-0.026
Mod	High	Mod	DV-adjusted	-0.082	-0.099	-0.116	-0.085	-0.033	-0.01	0.007
			E-adjusted	-0.17	-0.159	-0.157	-0.125	-0.081	-0.063	-0.05
			Naive	-0.193	-0.179	-0.173	-0.141	-0.098	-0.081	-0.073
Mod	High	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124

Table SM 7: Average bias of the estimators over time when the S_i's and inter-visit times follow a Weibull distribution – High conversion probability

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
High	Low	Low	DV-adjusted	0.081	0.097	0.088	0.04	0.031	0.02	0.017
			E-adjusted	-0.044	-0.031	-0.015	0.004	0.009	0.021	0.025
			Naive	-0.099	-0.078	-0.066	-0.048	-0.034	-0.028	-0.025
High	Low	Mod	DV-adjusted	-0.085	-0.099	-0.116	-0.086	-0.033	-0.01	0.007
			E-adjusted	-0.172	-0.16	-0.157	-0.125	-0.08	-0.063	-0.05
			Naive	-0.194	-0.179	-0.173	-0.141	-0.098	-0.081	-0.073
High	Low	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124
High	Mod	Low	DV-adjusted	0.081	0.097	0.088	0.04	0.031	0.019	0.016
			E-adjusted	-0.044	-0.031	-0.015	0.004	0.009	0.021	0.025
			Naive	-0.099	-0.078	-0.066	-0.048	-0.034	-0.028	-0.025
High	Mod	Mod	DV-adjusted	-0.083	-0.099	-0.117	-0.087	-0.034	-0.01	0.007
			E-adjusted	-0.171	-0.16	-0.158	-0.126	-0.081	-0.063	-0.05
			Naive	-0.193	-0.179	-0.173	-0.142	-0.099	-0.082	-0.073
High	Mod	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124
High	High	Low	DV-adjusted	0.081	0.097	0.088	0.04	0.031	0.02	0.017
			E-adjusted	-0.044	-0.031	-0.015	0.004	0.009	0.021	0.025
			Naive	-0.099	-0.078	-0.066	-0.048	-0.034	-0.028	-0.025
High	High	Mod	DV-adjusted	-0.084	-0.099	-0.116	-0.086	-0.033	-0.01	0.007
			E-adjusted	-0.171	-0.16	-0.158	-0.125	-0.081	-0.063	-0.05
			Naive	-0.194	-0.179	-0.173	-0.142	-0.099	-0.082	-0.073
High	High	High	DV-adjusted	0.011	-0.121	-0.184	-0.203	-0.131	-0.095	-0.076
			E-adjusted	-0.194	-0.235	-0.25	-0.24	-0.166	-0.132	-0.113
			Naive	-0.215	-0.249	-0.26	-0.248	-0.176	-0.143	-0.124

Table SM 8: SE of the estimators over time when the S_i's and inter-visit times follow an exponential distribution – Low conversion probability.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Low	Low	Low	True-DV-adjusted	0.134	0.109	0.084	0.061	0.042	0.047	0.031
			DV-adjusted	0.132	0.106	0.08	0.057	0.043	0.05	0.031
			E-adjusted	0.106	0.085	0.065	0.047	0.032	0.037	0.024
Low	Low	Mod	True-DV-adjusted	0.119	0.11	0.101	0.096	0.07	0.062	0.05
			DV-adjusted	0.123	0.104	0.09	0.082	0.06	0.053	0.043
			E-adjusted	0.09	0.087	0.08	0.075	0.053	0.047	0.038
Low	Low	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028
Low	Mod	Low	True-DV-adjusted	0.138	0.111	0.085	0.062	0.043	0.049	0.031
			DV-adjusted	0.135	0.108	0.081	0.058	0.044	0.051	0.031
			E-adjusted	0.109	0.087	0.066	0.048	0.033	0.039	0.024
Low	Mod	Mod	True-DV-adjusted	0.119	0.11	0.101	0.096	0.069	0.062	0.05
			DV-adjusted	0.123	0.104	0.09	0.082	0.059	0.053	0.043
			E-adjusted	0.09	0.087	0.08	0.075	0.053	0.047	0.038
Low	Mod	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028
Low	High	Low	True-DV-adjusted	0.138	0.111	0.084	0.061	0.04	0.049	0.031
			DV-adjusted	0.135	0.108	0.079	0.057	0.041	0.051	0.031
			E-adjusted	0.11	0.087	0.065	0.047	0.031	0.039	0.024
Low	High	Mod	True-DV-adjusted	0.119	0.11	0.101	0.096	0.069	0.062	0.05
			DV-adjusted	0.123	0.104	0.09	0.082	0.059	0.053	0.043
			E-adjusted	0.09	0.087	0.08	0.075	0.053	0.047	0.038
Low	High	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028

Table SM 9: SE of the estimators over time when the S_i's and inter-visit times follow an exponential distribution – Moderate conversion probability.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Mod	Low	Low	True-DV-adjusted	0.138	0.111	0.084	0.061	0.04	0.049	0.031
			DV-adjusted	0.135	0.108	0.079	0.057	0.041	0.051	0.031
			E-adjusted	0.11	0.087	0.065	0.047	0.031	0.039	0.024
Mod	Low	Mod	True-DV-adjusted	0.119	0.11	0.101	0.096	0.069	0.062	0.05
			DV-adjusted	0.123	0.104	0.09	0.082	0.059	0.053	0.043
			E-adjusted	0.09	0.087	0.081	0.075	0.053	0.047	0.038
Mod	Low	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028
Mod	Mod	Low	True-DV-adjusted	0.134	0.109	0.084	0.061	0.042	0.047	0.031
			DV-adjusted	0.132	0.106	0.08	0.057	0.043	0.05	0.031
			E-adjusted	0.106	0.085	0.065	0.047	0.032	0.037	0.024
Mod	Mod	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028
Mod	High	Low	True-DV-adjusted	0.135	0.109	0.084	0.061	0.043	0.048	0.031
			DV-adjusted	0.133	0.106	0.08	0.057	0.044	0.051	0.031
			E-adjusted	0.107	0.085	0.066	0.047	0.033	0.038	0.024
Mod	High	Mod	True-DV-adjusted	0.119	0.108	0.101	0.097	0.07	0.063	0.051
			DV-adjusted	0.122	0.102	0.09	0.083	0.06	0.054	0.044
			E-adjusted	0.09	0.085	0.081	0.076	0.054	0.047	0.038
Mod	High	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028

10

Table SM 10: SE of the estimators over time when the S_i's and inter-visit times follow an exponential distribution – High conversion probability.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
High	Low	Low	True-DV-adjusted	0.136	0.112	0.083	0.06	0.039	0.05	0.031
			DV-adjusted	0.133	0.109	0.079	0.057	0.039	0.051	0.03
			E-adjusted	0.108	0.088	0.065	0.047	0.03	0.039	0.024
High	Low	Mod	True-DV-adjusted	0.12	0.108	0.101	0.097	0.07	0.063	0.051
			DV-adjusted	0.123	0.102	0.09	0.083	0.06	0.054	0.044
			E-adjusted	0.091	0.085	0.08	0.076	0.054	0.047	0.038
High	Low	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028
High	Mod	Low	True-DV-adjusted	0.136	0.112	0.083	0.06	0.039	0.05	0.031
			DV-adjusted	0.133	0.109	0.079	0.057	0.039	0.051	0.03
			E-adjusted	0.108	0.088	0.065	0.047	0.03	0.039	0.024
High	Mod	Mod	True-DV-adjusted	0.12	0.108	0.101	0.097	0.07	0.063	0.051
			DV-adjusted	0.123	0.102	0.09	0.083	0.06	0.054	0.044
			E-adjusted	0.091	0.085	0.08	0.076	0.053	0.047	0.038
High	Mod	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028
High	High	Low	True-DV-adjusted	0.136	0.112	0.083	0.06	0.039	0.05	0.031
			DV-adjusted	0.133	0.109	0.079	0.057	0.039	0.051	0.03
			E-adjusted	0.108	0.088	0.065	0.047	0.03	0.039	0.024
High	High	Mod	True-DV-adjusted	0.12	0.108	0.101	0.097	0.07	0.063	0.051
			DV-adjusted	0.123	0.102	0.09	0.083	0.06	0.054	0.044
			E-adjusted	0.091	0.085	0.08	0.076	0.053	0.047	0.038
High	High	High	True-DV-adjusted	0.08	0.125	0.087	0.072	0.056	0.046	0.038
			DV-adjusted	0.094	0.142	0.093	0.069	0.048	0.038	0.031
			E-adjusted	0.048	0.083	0.065	0.059	0.044	0.035	0.028

Table SM 11: SE of the estimators over time when the S_i's and inter-visit times follow a Weibull distribution – Low conversion probability.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Low	Low	Low	DV-adjusted	0.143	0.118	0.099	0.077	0.06	0.054	0.047
			E-adjusted	0.121	0.102	0.087	0.068	0.049	0.044	0.038
Low	Low	Mod	DV-adjusted	0.141	0.119	0.107	0.092	0.063	0.059	0.049
			E-adjusted	0.111	0.105	0.099	0.089	0.06	0.056	0.046
Low	Low	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034
Low	Mod	Low	DV-adjusted	0.141	0.117	0.098	0.076	0.06	0.055	0.045
			E-adjusted	0.119	0.101	0.086	0.068	0.049	0.045	0.037
Low	Mod	Mod	DV-adjusted	0.14	0.119	0.107	0.092	0.063	0.059	0.049
			E-adjusted	0.111	0.105	0.099	0.089	0.06	0.056	0.046
Low	Mod	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034
Low	High	Low	DV-adjusted	0.141	0.117	0.097	0.075	0.06	0.055	0.045
			E-adjusted	0.119	0.101	0.085	0.067	0.049	0.045	0.037
Low	High	Mod	DV-adjusted	0.14	0.119	0.107	0.092	0.063	0.059	0.049
			E-adjusted	0.111	0.105	0.099	0.089	0.06	0.056	0.046
Low	High	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034

3.2 Study 2

Tables SM11–SM13 show the SE of the estimators in study 2 where we generate Weibull distributed the S_i's and inter-visit times.

Table SM 12: SE of the estimators over time when the S_i's and inter-visit times follow a Weibull distribution – Moderate conversion probability.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
Mod	Low	Low	DV-adjusted	0.141	0.117	0.097	0.075	0.06	0.055	0.045
			E-adjusted	0.119	0.101	0.085	0.067	0.049	0.045	0.037
Mod	Low	Mod	DV-adjusted	0.14	0.119	0.107	0.092	0.063	0.059	0.049
			E-adjusted	0.111	0.105	0.099	0.089	0.06	0.056	0.046
Mod	Low	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034
Mod	Mod	Low	DV-adjusted	0.143	0.118	0.099	0.077	0.06	0.054	0.047
			E-adjusted	0.121	0.102	0.087	0.068	0.049	0.044	0.038
Mod	Mod	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034
Mod	High	Low	DV-adjusted	0.142	0.119	0.099	0.077	0.06	0.054	0.047
			E-adjusted	0.12	0.103	0.087	0.069	0.049	0.045	0.038
Mod	High	Mod	DV-adjusted	0.142	0.121	0.108	0.093	0.063	0.059	0.049
			E-adjusted	0.112	0.106	0.1	0.09	0.061	0.056	0.046
Mod	High	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034

Table SM 13: SE of the estimators over time when the S_i's and inter-visit times follow a Weibull distribution – High conversion probability.

Conversion probability	Number of visits	Delay time	Estimators	Days since the first click						
				10	15	21	30	39	51	60
High	Low	Low	DV-adjusted	0.14	0.117	0.096	0.075	0.061	0.053	0.046
			E-adjusted	0.119	0.101	0.084	0.067	0.05	0.043	0.037
High	Low	Mod	DV-adjusted	0.142	0.12	0.108	0.093	0.063	0.059	0.049
			E-adjusted	0.112	0.106	0.1	0.09	0.061	0.056	0.046
High	Low	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034
High	Mod	Low	DV-adjusted	0.141	0.117	0.096	0.075	0.061	0.053	0.046
			E-adjusted	0.119	0.101	0.084	0.067	0.05	0.043	0.037
High	Mod	Mod	DV-adjusted	0.143	0.121	0.108	0.093	0.063	0.059	0.049
			E-adjusted	0.113	0.106	0.1	0.09	0.06	0.056	0.046
High	Mod	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034
High	High	Low	DV-adjusted	0.14	0.117	0.096	0.075	0.061	0.053	0.046
			E-adjusted	0.119	0.101	0.084	0.067	0.05	0.043	0.037
High	High	Mod	DV-adjusted	0.143	0.121	0.108	0.093	0.063	0.059	0.049
			E-adjusted	0.113	0.106	0.1	0.09	0.061	0.056	0.047
High	High	High	DV-adjusted	0.125	0.135	0.098	0.08	0.054	0.043	0.035
			E-adjusted	0.075	0.088	0.072	0.069	0.051	0.041	0.034