CTV Co-viewing Rate Estimation Using Online Surveys

Rachel Fan, Wei Liu, and Jon Vaver

Google Inc.

Abstract

Co-viewing refers to the situation in which multiple people share the experience of watching video content and ads in the same room and at the same time. In this paper, we use online surveys to measure the co-viewing rate for YouTube videos that are watched on a TV screen. These simple one-question surveys are designed to optimize response accuracy. Our analysis of survey results identifies variations in co-viewing rate with respect to important factors that include the demographic group (age/gender) of the primary viewer, time of day, and the genre of the video content. Additionally, we fit a model based on these covariates to predict the co-viewing rate for ad impressions that are not directly informed by a survey response. We present results from a case study and show how co-viewing changes across these covariates.

1 Background

Connected TV (cTV) refers to a TV set connected to the internet either as a SmartTV or via an external device such as chromecast, digital media player or gaming console. In 2021, nearly 83% of households will have at least one connected TV [1]. With viewership surging, cTV is now the fastest growing platform for watching YouTube videos. In December 2020, more than 120 million people in the U.S. streamed YouTube or YouTube TV on their TV screen [2] . The rapid growth of this platform makes it important to better understand how it is being used.

When a video is played on a cTV screen, more than one person may be watching the content at the same time. This living room "co-viewing" behavior is prevalent on TV screens when people consume live linear TV[1] content. According to a recent study conducted by Neilsen with Google, YouTube is watched by multiple viewers (ages 18 and above) 26% of the time compared to 22% for linear TV [1].

Having the ability to measure the co-viewing behavior of users on cTV makes it possible to measure advertising reach on cTV in a manner that is comparable to that of traditional TV advertising. There are multiple approaches to measure the co-viewing rate. A general standard is to build a high-quality panel that is representative of the target audience. For households recruited to participate in a cTV panel, their cTV devices are electronically monitored by a meter. Panelists are asked to check-in every time they watch the cTV device and this makes it possible to observe their co-viewing behaviors. Panels also provide reliable profile information on their members that can be used for deeper analysis. On the other hand, recruiting panelists to participate in a high quality representative panel is expensive and usually takes many months. Maintaining such a panel is also challenging due to the possibility that panelists will become less compliant over time. Finally, representative panels usually include thousands, or perhaps tens of thousands, of households. This limits the ability to perform granular analyses with multiple dimensions.

In this study, we explore an alternative approach that deploys online surveys to collect information about co-viewing behavior. In particular, we send out surveys to active cTV users to ask them how many people are watching at the moment when an ad would otherwise be served during a YouTube viewing session. Collecting a large number of survey responses makes it possible to develop a model to predict cTV co-viewing rate from covariates including demographic groups, time of day slices and genre, as described below.

[1]Linear TV is the traditional television system in which the viewer can only watch a scheduled TV program on a particular channel that offers the content.

2 Surveys

2.1 Survey questions and answer choices

The survey used to collect co-viewing data has a single question. The wording of the question is "Including yourself, how many people are watching this TV right now?" There are 5 options to choose from: "1", "2", "3", "4 or more" or "Prefer not to answer". This requires a single-choice response, so users can only choose one option. The order of the options is fixed because there is a logical order in the choices. We use both the original order and the reversed order in our survey to mitigate possible biases from the order of options; users are randomly assigned to one of the orders. There is a "SKIP SURVEY" button at the lower right corner so users can choose not to answer the survey.

This 1-question survey creates a simple and consistent user experience, and the skip button reduces the incentive to randomly select an answer so that a thoughtful and accurate response is elicited.

2.2 Survey setup

Surveys appear before users begin to watch a video on a cTV screen. They are shown in place of an ad that would otherwise appear at the same time. Surveys are shown over the entire screen area of the video player. If left unattended, surveys will disappear after 30 seconds and the video will begin to play. Users can either answer the survey, skip the survey or wait for the survey to time out. Surveys are served to a random sample of 7-day active YouTube cTV users[2]. Each user will not receive the survey more than once every 35 days. The vast reach of YouTube provides a wide user base to survey. The surveys used for this analysis were collected between October 2020 and April 2021 in the US. This collection generated over one million survey responses globally.

2.3 Survey quality control

Response accuracy is crucial for an accurate estimate of co-viewing rate. We use signals to filter out responses that are unlikely to be thoughtful. For example, if a user responds too quickly to generate a trustworthy response, this response is discarded because it is less likely to be accurate.

As in any survey, solicitation biases and response biases are inevitable. To be solicited, a user in the solicitation pool must watch a YouTube video on a cTV device, hence more active cTV users have a higher chance of being solicited. Response bias comes from user choices. If users do not want to answer the survey they can skip the survey or wait for the survey to disappear. Also, users might be more likely to answer if there are fewer (or more) coviewers. These factors may make the results from survey respondents different from the YouTube cTV population. To reduce solicitation and response biases, our methodology calibrates survey results with YouTube log data, as described below.

3 Co-viewing rate definition

How we define the co-viewing rate will substantially affect our collection of data, the calculations used to generate results, and how these results should be applied. Important factors include the treatment of presence and when the coviewing rate is measured.

3.1 Co-viewing and presence

User presence refers to the possibility that no viewer may be present when a video, or ad, is shown. A TV may be left unattended while a video or ad plays.

In our survey-based results, the co-viewing rate is defined to be the number of viewers given that at least one viewer is present. This is a necessary assumption with a survey-based approach because at least one viewer must be present to respond to the survey. With this definition, co-viewing rate will always be greater than or equal to 1 at both the impression level and aggregated level. (A definition that combines presence and co-viewing rate would always be greater than or equal to zero.) In addition to accommodating survey-based measurement, this definition also has the advantage of separating the measurement of presence from the measurement of co-viewing rate. This simplifies both the measurement and modeling of presence and co-viewing rate. One measurement/model can be used to determine if at least one viewer is present and a separate measurement/model can be used to determine the number of co-viewers. This separation also simplifies the demographic modeling of viewers, since such a model is not applicable without presence.

3.2 Temporal placement of co-viewing measurement

Generally speaking, there are many possible choices for when and how to measure the co-viewing rate for

[2]Google users refer to cookies. One person may have multiple Google accounts, hence multiple cookies. And multiple people may share one cookie.

cTV. These include:

- position in the video: the number of viewers might be measured at the beginning, middle, or end of a video playback
- level of resolution: the number of viewers might be measured at each second/minute of a video; and a co-viewing rate might be generated by computing the average, minimum, or maximum number of viewers across a single video playback or a session with multiple video playbacks
- content-based or ad-based measurement: the number of viewers may be measured during the video content or during an ad that plays immediately before, during, or after the video content

In this analysis, our survey-based measurement of co-viewing rate takes place at the point of an ad impression that occurs immediately before the beginning of a video playback. The measurement occurs at this point because surveys directly replace a YouTube ad that precedes a video playback. The dynamics of co-viewing are not captured with this definition of co-viewing. For example, more people might join in watching the video after it has started, and some people who are present at the beginning of the video may not watch the entire video. On the other hand, measuring the co-viewing rate at the point of ad serving is the most relevant choice for advertisers. Surveys that replace ads in the middle and at the end of the video will be studied in future analyses.

4 Co-viewing rate estimation

The cTV co-viewing rate is estimated using eligible survey responses from the survey. The "None of the above" responses are removed from the data as are responses that were submitted too quickly. To obtain an unbiased result, responses are weighted using a calibration process to match the characteristics of survey respondents to those of the YouTube cTV population identified from YouTube logs data. Finally, these data are used to estimate an overall coviewing rate and to build a model that predicts co-viewing rate at the impression level using demographic information, time of day, and video genres.

4.1 Overall co-viewing rate

The overall co-viewing rate provides an overall picture of co-viewing behavior and its impact on ad reach. It is also the foundation for estimating co-viewing at the impression level, which is described in the next section.

Suppose we have n survey responses. For $i \in \{1, ..., n\}$, define

- y_i : survey response i, i.e. the number of co-viewers in response i, $y_i \in \{1, 2, 3, 4+\}$
- x_{1i}: the demographic (age/gender) group of the respondent for response i
- x_{2i}: the time of day/week slice for response i
- x_{3i}: the genre associated with video content for response i

The smallest value of y_i is 1 because at least one person has to be present to answer the survey and this person is the "main user". The largest number is "4+" representing "4 or more" in the survey options. In our calculation, we replace "4+" with "4", therefore our result is a conservative estimate of the co-viewing rate. About 8% of all valid survey responses in the US chose "4+"; if among them 5% has 4 people watching and 3% has 5 people watching, the overall co-viewing rate is underestimated by about 0.03. An intuitive estimate of overall co-viewing rate is the average of all y_i, i.e. $\frac{\Sigma_i y_i}{n}$. However, this simple estimate may be biased due to solicitation bias and response bias, as discussed in the Section Survey quality control.

We can improve this estimate with post-sample calibration [3, 5], which uses weights to match the covariates distribution of the respondents to that of the YouTube cTV population. For example, if older age groups are more likely to respond to the survey and older age groups tend to have high co-viewing rates, the simple estimate defined above will overestimate the co-viewing rate. This bias is mitigated by giving smaller weights to older respondents and higher weights to younger respondents so that the weighted sample is more representative of the YouTube cTV population. We apply the empirical calibration algorithm with quadratic loss [6] and calibrate against the distribution of viewers in the YouTube cTV population across demographic groups, time of day slices, and genre. The overall co-viewing rate is calculated using the resulting weights

$$\overline{y} = \frac{\Sigma_i w_i y_i}{\Sigma w_i} \tag{1}$$

where w_i is the weight assigned to response i from calibration.

4.2 Co-viewing factor at the covariate level

Co-viewing rate for cTV at the impression level depends on multiple covariates, including the demographic group, time of day slice and video genre. For instance, we observe that females have a higher co-viewing rate than males, and the co-viewing rate is higher when people are watching movie videos versus fitness videos. Therefore, it is useful to assess variations in co-viewing rate at the covariate level. However, there can be confounding effects between different covariates. For example, if we observe that (1) most respondents during prime time are females and (2) both prime time respondents and females have higher than average co-viewing rates, it is difficult to tell whether the high co-viewing rate for females is due to gender or because they are more likely to respond during prime time. To remove such confounding effects, we perform a calibration within slices of each covariate to ensure that the distributions of other covariates within this slice match a benchmark distribution before calculating this slice level co-viewing rate.

Below we illustrate the co-viewing factor calculation for demographic groups. Let D be the set of all possible values of demographic groups, i.e. $D = \{\text{F18-24}, \text{F25-34}, \ldots, \text{M18-24}, \text{M25-34}, ...\}$, where F18-24 corresponds to females between 18 and 24 years of age and M25-34 corresponds to males between 25 and 34 years of age, etc. For each value $d \in D$ (e.g. $d = \text{F18-24}$), let S_d be the set of all impressions with value d, i.e. $S_d = \{i \in \{1, ..., n\} | x_{1i} = d\}$. Within set S_d we perform empirical calibration to match the time of day distribution and genre distribution to that of the YouTube cTV population and accordingly assign a weight w_i^d for each observation i in S_d. Then, the calibrated co-viewing rate of S_d is

$$\overline{y_d} = \frac{\Sigma_{i \in S_d} w_i^d y_i}{\Sigma_{i \in S_d} w_i^d} \tag{2}$$

The Relative Co-viewing Factor (RCF) for d is defined as

$$E_d = \frac{\overline{y_d} - 1}{\overline{y} - 1} \tag{3}$$

where $\overline{y}$ is the overall co-viewing rate defined in (1). By subtracting 1 in both numerator and denominator, we remove the count of the main user who, by definition, must be present. Hence, the relative co-viewing factor is focused on the additional viewers (that exclude self) present in set S_d.

Similarly, let T be the set of all possible time of day slices and G be the set of all possible genres. The co-viewing factor can be calculated for each value $t \in T$ or each $g \in G$ in the same way.

4.3 A case study on co-viewing factor in the US

In this section we show the variations of co-viewing rate at covariate level from the survey responses collected from October 2020 to April 2021 in the US. We apply the methodology described above and calculate the relative co-viewing factor (RCF) by demographic group, time of day slice, and genre group. RCFs are not the absolute co-viewing rates, instead, they are relative to the overall co-viewing rate when excluding self.

4.3.1 Relative co-viewing factor by demographics

User age and gender associated with survey responses are estimated in Google logs data; the qualities of the demographic signals are discussed in this paper [4]. Figure 1 shows the relative co-viewing factor by demographics. The respondent's demographic group is shown on the horizontal axis and the vertical axis indicates the RCF associated with each demographic group. For example, when the respondent is in the demographic group F25-34 (i.e. females between 25 and 34 years old), the co-viewing rate minus 1 is, on average, 20% higher than the overall cTV co-viewing rate minus 1. The "Unknown" group contains respondents without known demographic information. The vertical bars in this plot indicate the 95% confidence intervals for the relative co-viewing factor. Results show that middle-aged users are more likely to watch videos on cTV together with their family members or friends. This is true for males and females. In contrast, younger viewers in the 18-24 age group and more senior viewers of 45+ have a higher tendency to watch alone. In addition, females have a slightly higher co-viewing rate than males in most age groups, indicating that females are more likely to watch with companions.

4.3.2 Relative co-viewing factor by time-of-day

Figure 2 demonstrates the relative co-viewing factor by time-of-day. The horizontal axis indicates the hour of day and the vertical axis indicates the value of RCF. Again, the vertical error bars show the 95% confidence intervals for the RCF. The blue line corresponds to the RCF for weekdays and the red line corresponds to weekends. There are many fewer responses in some dayparts, especially those in the overnight hours. As a result, the 24-hour day is partitioned into 9 distinct slices with non-uniform width to ensure a reasonable sample size in

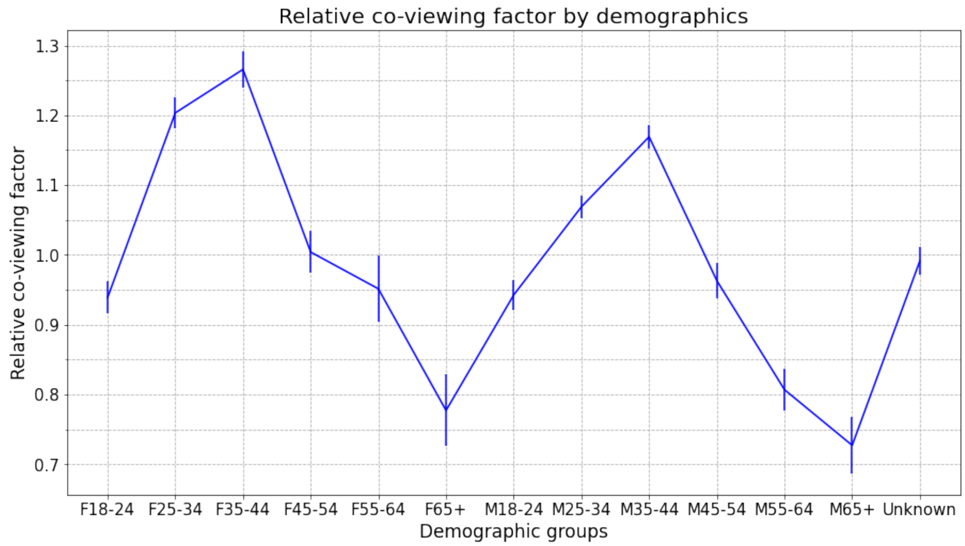

Figure 1: Relative co-viewing factor by demographics. The horizontal axis is the demographic group. "F" stands for female and "M" stands for male. "F18-24" corresponds to females between 18 and 24 years of age. The "Unknown" group contains users without known demographic information. The vertical axis is the relative co-viewing factor. The vertical bars indicate the 95% confidence intervals for the relative co-viewing factor.

each group. Weekends have a higher co-viewing rate than weekdays for each time slice. On both weekdays and weekends, the co-viewing rates are relatively stable from morning to early afternoon (8am-3pm) and have a big increase starting at 4pm. Prime time (6pm-9pm) has the highest co-viewing rate, and nighttime (2am-7am) has the lowest co-viewing rate.

4.3.3 Relative co-viewing factor by genre

We use a Google video classification to label video genres. This is a hierarchical multi-label classification in a tree structure, where a parent node of a large group such as "Media & Entertainment" can be divided into several sub-groups including "Music & Audio" and "Movies", which can be further divided into smaller sub-groups. There are hundreds of terminal nodes in the classification tree. Some of them have small traffic volume and very few survey responses, therefore, a bottom-up aggregation approach is applied to group the genre labels. We start from genres in the terminal node; if there are very few responses in the terminal node, we roll up the corresponding survey responses to the parent node and continue to do so until there are at least $2,000$ [3] responses in the node or we reach the root node. In the end, we have about fifty aggregated genre groups. Figure 3 shows the relative co-viewing factor of selected genre groups. The horizontal axis is the genre group and vertical axis is the relative co-viewing factor. It is clear from this plot that different genre groups can have very different co-viewing rates. Videos related to movies and music tend to have higher co-viewing rates while videos related to fitness and electronics generally have lower co-viewing rates.

5 Co-viewing rate prediction

Survey responses are collected at the level of individual ad impression. It is useful to be able to find the co-viewing rate for aggregations of impressions across a campaign, an advertiser, or a vertical. Therefore, we build a model to predict the co-viewing rate using important signals for each impression which can be aggregated for various use cases.

5.1 A three-predictor multiplicative model

We include demographic groups, time-of-day slices, and video genres as predictors in the model. Define S_{dtg} as the set of all the impressions with demographic group d, time of day slice t, and genre g, i.e. $S_{dtg} = \{i \in 1, ..., n | x_{1i} = d, x_{2i} = t, x_{3i} = g\}$. Our

[3]The threshold 2,000 was used to create a reasonable number of genres with diverse co-viewing rates. Lowering the threshold to include additional genre sub-groups does not provide additional resolution of co-viewing rates. For example, co-viewing rates for sub-groups of "Music & Audio" are similar to that of the parent node.

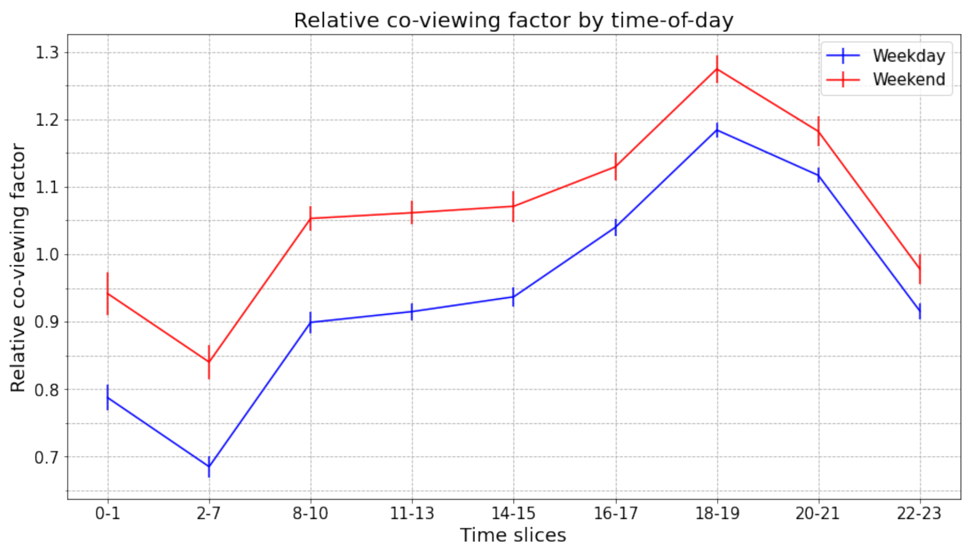

Figure 2: Relative co-viewing factor by time-of-day. The horizontal axis is the time slice. The vertical axis is the relative co-viewing factor. The vertical bars indicate the 95% confidence intervals for the relative co-viewing factor. The blue line represents the RCF on weekdays and the red line represents the weekends.

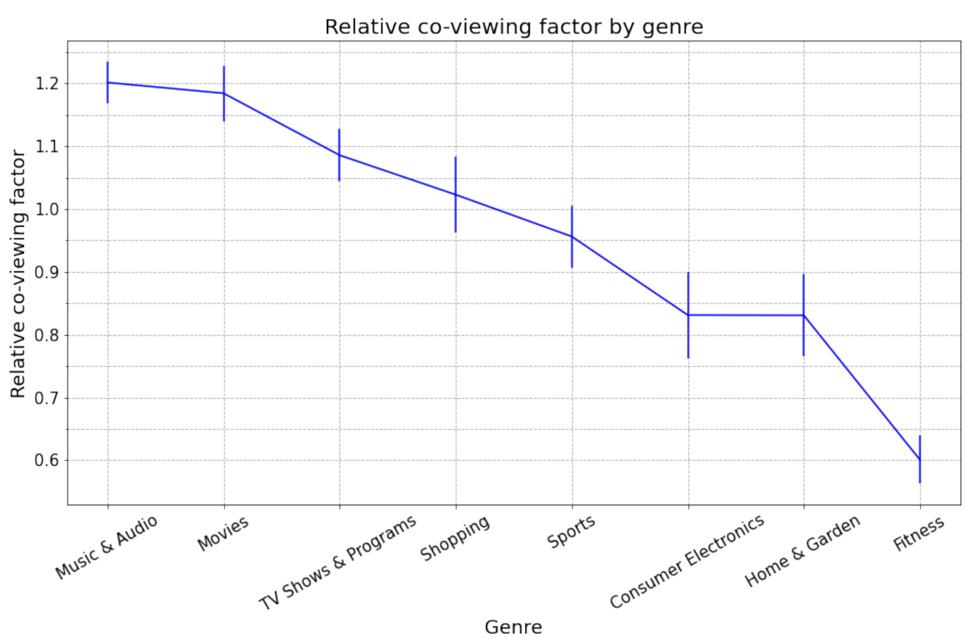

Figure 3: Relative co-viewing factor in selected genre groups. The horizontal axis is the genre group. The vertical axis is the relative co-viewing factor. The vertical bars indicate the 95% confidence intervals of the relative co-viewing factor.

goal is to predict the co-viewing rate for impressions in S_{dtg}. A straightforward approach is use all of the survey responses from S_{dtg} to calculate an average co-viewing rate directly. However, this approach suffers from data sparsity, especially for those sets with low traffic and a small number of responses. For example, with 13 demographic groups, 48 time of day slices, and hundreds of genre groups, there are about $60,000$ unique combinations of these three covariates. Having at least 100 responses per covariate combination requires at least 6 million survey responses. In practice, many more responses are needed to achieve the target number of responses considering that some combinations will have a very low volume of traffic. Therefore, we choose a modeling approach as a practical alternative for dealing with data sparsity.

A multiplicative model is used to predict the co-viewing rate in set S_{dtg}. Using equation (3) we first compute the co-viewing factors E_d, E_t, and E_g for demographic group d, time of day slice t, and genre g. The joint co-viewing factor of S_{dtg} is defined as the product of these separately generated individual co-viewing factors:

$$E_{dtg} = E_d * E_t * E_g \tag{4}$$

The predicted co-viewing rate for set S_{dtg} is:

$$\hat{y}_{dtg} = E_{dtg} * (\bar{y} - 1) + 1 \tag{5}$$

where $\bar{y}$ is the overall co-viewing rate from the equation (1).

This multiplicative model was chosen over many other models that were considered because it has a good prediction performance, better interpretability, and it generates predictions that have the same range as the co-viewing rate $[1, \infty]$ (See Appendix A for details.)

5.2 Population traffic normalization

We introduce a normalization factor r to further align the results with the traffic patterns of the YouTube cTV population. Let $D \times T \times G$ denote the Cartesian product of sets D, T and G, which consists of all triplets $(d, t, g)$for which $d \in D$, $t \in T$ and $g \in G$. Define

$$r = \frac{1}{\Sigma_{(d,t,g) \in D \times T \times G} p_{dtg} E_{dtg}} \tag{6}$$

where E_{dtg} is the joint co-viewing factor of impressions in set S_{dtg} as defined in the equation (4) and p_{dtg} is the percentage of impressions in set S_{dtg} relative to the entire YouTube cTV population. The normalized joint co-viewing factor of set S_{dtg} is

$$E^N_{dtg} = r * E_{dtg} \tag{7}$$

The normalized co-viewing rate for set S_{dtg} is:

$$\hat{y}^N_{dtg} = E^N_{dtg} * (\bar{y} - 1) + 1 \tag{8}$$

Including this normalization factor ensures that the overall co-viewing rate remains equal to $\bar{y}$ after the multiplicative co-viewing rate model is applied to all impressions. This is necessary because we assume that the effects from different covariates are independent and multiplicative. If the real world scenario deviates from this assumption, the average co-viewing rate of the entire YouTube cTV traffic will be different from the overall co-viewing rate $\bar{y}$ when aggregating over all impressions.

5.3 Comparison of predicted and observed co-viewing factors

Let the Covariate Co-viewing Factor (CCF) be the ratio of the co-viewing rate EXCLUDING self predicted for a particular combination of covariates (demographic, time of day, genre) normalized by the overall co-viewing rate, i.e.

$$CCF = \frac{\hat{y}^N_{dtg} - 1}{\bar{y} - 1} \tag{9}$$

We compare the predicted CCF to the observed CCF for covariate combinations with more than 50 responses. There are about $1,400$ such combinations and the median response count in these groups is 99. Fig. 4 shows a scatter plot of the observed CCF versus the predicted CCF for these covariate combinations. The correlation coefficient between the observed CCF and predicted CCF is 0.82. The difference between the average predicted CCF and the average observed CCF (weighted by the number of respondents) is 0.05. Predictions from this multiplicative model tend to overestimate when the co-viewing rate is low and underestimate when the co-viewing rate is high.

6 Discussion

Accounting for the co-viewing rate on cTV is essential for an accurate estimation of reach in digital ads and for comparing this reach to that of traditional television. Real-time online survey provides a cost-effective and time-efficient approach for calculating an empirical estimate of cTV co-viewing rate. In this paper, we describe an approach for measuring the cTV co-viewing rate using large-scale online surveys. This method generates the overall co-viewing rate, it demonstrates how covariates influence the co-viewing rate, and it predicts the co-viewing rate at

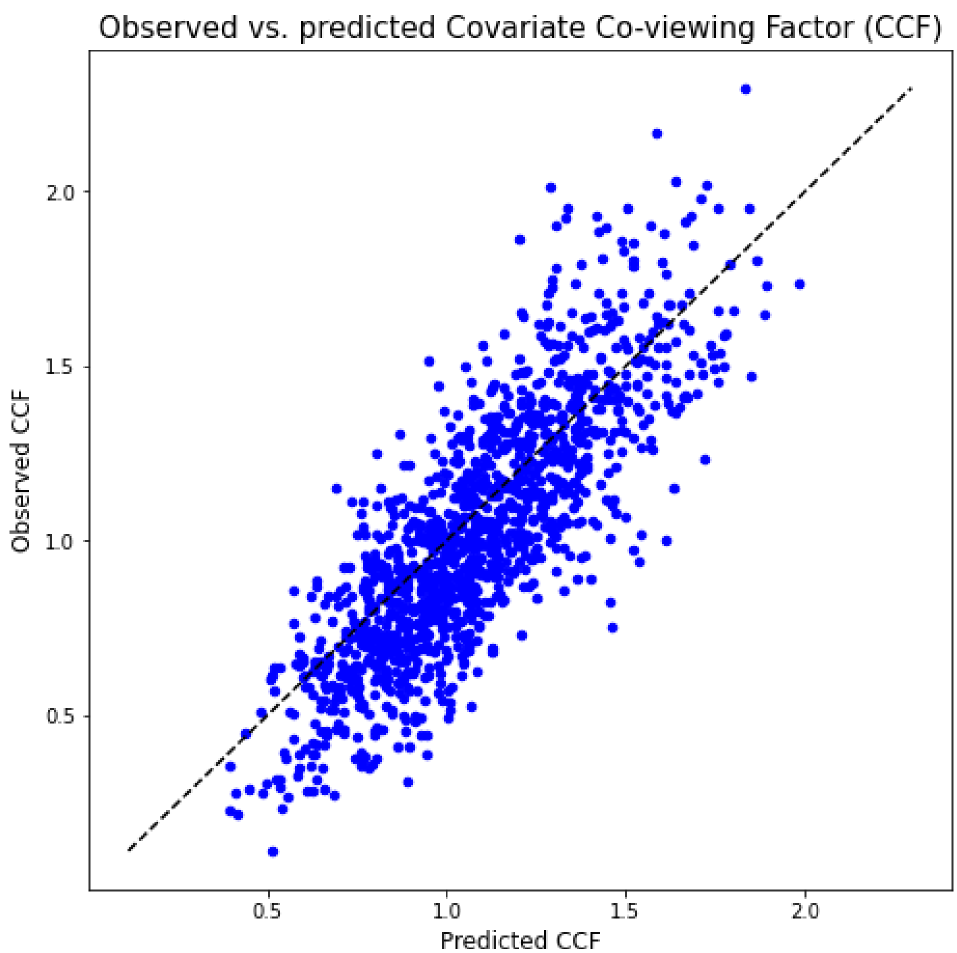

Figure 4: Predicted CCF vs. observed CCF for many combinations of (demographic, time of day, gender). The black dashed line is the identity line.

the impression level based on demographic group, time of day slice, and video genre. We showcase the results from a case study which illustrates the variation in cTV co-viewing rate with respect to these covariates of interest.

It is desirable to compare the empirical estimates of cTV co-viewing rate based on surveys with other "ground truth" measurements, such as those obtained from representative panels. There are ongoing efforts at Google to make such comparisons and to determine how these sources of information might be combined to further improve co-viewing rate measurement. In addition, independent measurement providers have their own panels and meters to measure cTV viewership. Future analysis efforts will also include comparisons of co-viewing results across platforms and methodologies. We hope this analysis will generate more research and discussions on cTV co-viewing rate measurement.

Acknowledgements

We would like to thank many colleagues at Google for their comments and feedback. In particular, we would like to acknowledge Jim Koehler, Jeremy Stone, Evgeny Skvortsov and Alex Johnston for their valuable discussions and suggestions.

Appendix A

Multiplicative model vs. linear regression model

In this paper, we describe a multiplicative model to predict the co-viewing rate of an ad impression given demographic group, time of day slice, and video genre. We also considered other predictive models such as linear regression and multinomial logistic regression before making this choice. Model performance was evaluated using the prediction root mean square error (RMSE) in 10-fold cross validation. Table A.1 lists the performance for a series of models, including a base model, models with one, two, or three predictors, and models that include interaction terms.

The base model assumes that every impression has the same predicted co-viewing rate as the overall co-viewing rate across survey responses calculated using Equation 1. Compared to this base model, the models with one predictor have improved RMSE in both in-sample fitting and out-of-sample prediction (see Table A.1). For models with two predictors, we evaluated multiplicative models and linear regression models with and without interactions. These two-predictor models have better performance than the single-predictor models and similar performance across different instances. Similarly, the three-predictors models outperform the two-predictor models. The final row of Table A.1 shows that the three-predictor model with a demo : genre interaction term has a higher degree of overfitting than the three-predictor multiplicative and linear regression models. As a result, these latter two models are favored over the other models in this table.

We continued the evaluation by examining co-viewing rate prediction at the covariate level. For both the three-predictor multiplicative and linear regression models, we compare the predicted co-viewing rate to the average observed co-viewing rate within groups of survey responses having the same demographic group, time of day slice, and video genre. Only groups with at least 20 responses were included in this analysis. More than 2500 groups of survey responses meet this condition. Figure A.1 contains a plot of the average observed co-viewing rate versus the co-viewing rate predicted by the three-predictor multiplicative model and three-predictor linear regression model. Predictions from both models tend to overestimate when the co-viewing rate is low and underestimate when the co-viewing rate is high. One caveat with the linear regression model is that for some combinations of covariates, the predicted co-viewing rate is below 1. This is inconsistent with our definition of co-viewing rate which assumes the presence of at least one viewer and requires the co-viewing rate to be greater than or equal to 1. Therefore, we choose the three-predictor multiplicative model over the three-predictor linear regression model.

Table A.1: Performance of different models to predict co-viewing rate

Model		Predictors	RMSE in in-sample fitting	RMSE in out-of-sample prediction
Base model		raw average	0.9344	0.9352
One-predictor models		demo	0.9299	0.9305
		time of day	0.9399	0.9305
		genre	0.9282	0.9283
Two-predictor models	multiplicative model	demo, time of day	0.9254	0.9259
		demo, genre	0.9237	0.9238
		time of day, genre	0.9241	0.9243
	linear regression	demo, time of day	0.9254	0.9256
		demo, genre	0.9237	0.9240
		time of day, genre	0.9242	0.9245
	linear regression with interactions	demo, time of day	0.9246	0.9259
		demo, genre	0.9212	0.9241
		time of day, genre	0.9218	0.9260
Three-predictor models	multiplicative model	demo, time of day, genre	0.9197	0.9200
	linear regression	demo, time of day, genre	0.9197	0.9201
	linear regression with (demo,genre) interaction	demo, time of day, genre	0.9172	0.9203

This table lists the root mean square error (RMSE) in in-sample fitting and out-of-sample prediction in a series of models, including the base model (which assumes every impression has the same co-viewing rate) and models with one, two or three predictors with and without interaction terms.

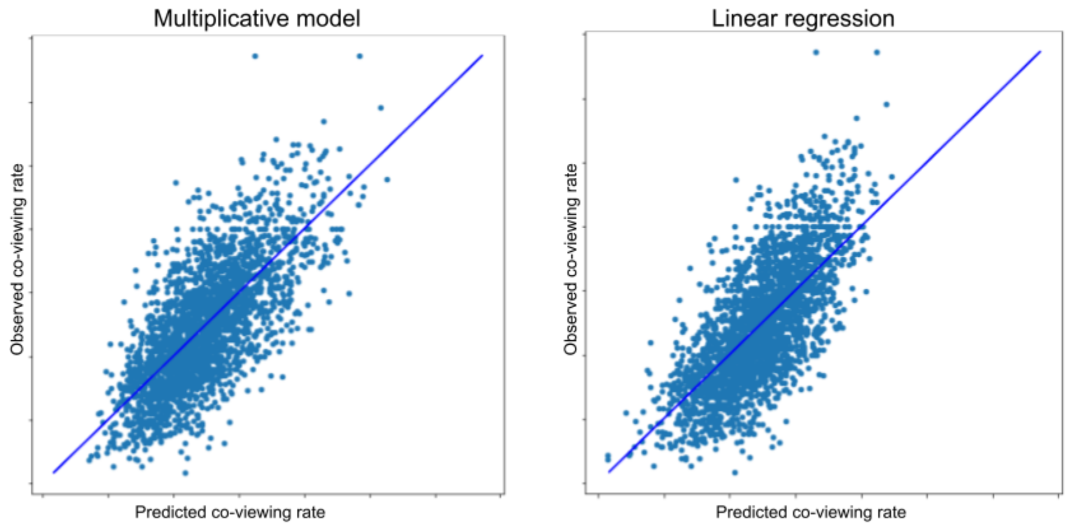

Figure A.1: Observed co-viewing rate vs. the co-viewing rate predicted by the three-predictor multiplicative and linear regression models. The blue line is the identity line.

References

[1] 3 ways top brands are using connected tv to increase reach and drive results. Technical report, Think with Google, April 2021.

[2] Accelerated by the pandemic, 3 trends that are redefining video. Technical report, Think with Google, March 2021.

[3] Jean-Claude Deville and Carl-Erik Särndal. Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87(418):376-382, 1992.

[4] Jim Koehler, Evgeny Skvortsov, and Wiesner Vos. A method for measuring online audiences. Technical report, Google Inc, 2013.

[5] Phillip S Kott. Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, 32(2):133, 2006.

[6] Xiaojing Wang, Jingang Miao, and Yunting Sun. A python library for empirical calibration. *arXiv preprint arXiv:1906.11920v2*, 2019.