ELEANOR MCDONNELL FEIT, PENGYUAN WANG, ERIC T. BRADLOW, and PETER S. FADER*

As firms collect greater amounts of data about their customers from an ever broader set of "touchpoints," a new set of methodological challenges arises. Companies often collect data from these various platforms at differing levels of aggregation, and it is not clear how to merge these data sources to draw meaningful inferences about customer-level behavior patterns. In this article, the authors provide a method that firms can use, based on readily available data, to gauge and monitor multiplatform media usage. The key innovation in the method is a Bayesian data-fusion approach that enables researchers to combine individual-level usage data (readily available for most digital platforms) with aggregated data on usage over time (typically available for traditional platforms). This method enables the authors to disentangle the intraday correlations between platforms (i.e., the usage of one platform vs. another on a given day) from longer-term correlations across users (i.e., heavy/light usage of multiple platforms over time). The authors conclude with a discussion of how this method can be used in a variety of marketing contexts for which data have become readily available, such as gauging the interplay between online and brick-and-mortar purchasing behavior.

Keywords: data fusion, Bayesian multivariate model, multiplatform behavior, media usage

Fusing Aggregate and Disaggregate Data with an Application to Multiplatform Media Consumption

*Eleanor McDonnell Feit is Executive Director, Wharton Customer Analytics Initiative, and Lecturer in Marketing, The Wharton School, University of Pennsylvania (e-mail: efeit@wharton.upenn.edu). Pengyuan Wang is a research scientist, Yahoo! Labs (e-mail: pengyuan@wharton.upenn.edu). Eric T. Bradlow is the K.P. Chao Professor of Marketing, Statistics, and Education, the Co-Director of the Wharton Customer Analytics Initiative, and Vice-Dean, Wharton Doctoral Programs (e-mail: ebradlow@wharton.upenn.edu), and Peter S. Fader is the Frances and Pei-Yuan Chia Professor of Marketing and the Co-Director of the Wharton Customer Analytics Initiative (e-mail: faderp@wharton.upenn.edu), The Wharton School, University of Pennsylvania. The authors recognize the support of the Wharton Customer Analytics Initiative and are grateful to the companies that provided the data for this study through this research center: ESPN provided disaggregate data on consumers' usage of digital media platforms, and The Nielsen Company provided aggregate data on television usage. Artie Bulgrin, Glenn Enoch, Jarvis Bowers, and Kim Weller of ESPN also provided helpful comments on an earlier version of this work. Joel Huber served as guest editor, and Carl Mela served as associate editor for this article.

When U.S. fans wanted to know what happened on a particular day of the 1990 Fédération Internacionale de Football Association (FIFA) World Cup tournament, they had few choices: they could see a final score on the evening news or read about it in a newspaper the following day. In sharp contrast, World Cup fans in 2010 could choose among numerous media platforms to follow every game in real time: they could follow continually updated coverage on a traditional website or on their mobile phones; they could watch every game live on television; or they could watch online streaming video, either live or recorded. This rapid proliferation of media delivery platforms is changing how media is consumed—not just for the World Cup but for all types of content. From a business perspective, the multitude of platforms provides both opportunities for greater media exposure and higher advertising reach (the "currency" of the media business) but also poses greater challenges for media companies trying to decide whether to invest in developing content for each new platform. To make these decisions, companies need tools to monitor how media consumers use the many available channels and the changing relationships between them.

Luckily, there is a wealth of behavioral data on how and when media users access content that can answer these questions. All the emerging media platforms—including websites, mobile devices, and online streaming video platforms—create a record of what each user has viewed, providing a rich opportunity for companies to investigate media consumption at the individual level. In many cases, firms can track a user's behavior across platforms. As media consumption migrates to these highly measurable platforms, companies regularly have the data necessary to constantly monitor how people use multiple platforms at the same time. However, data on the viewing of traditional media, particularly television (the most widely used medium), remain scant. Globally, only a small fraction of households have transitioned to measurable television systems, and even if a cable or satellite provider recorded television usage for an individual household, technology differences between television and the Internet make it difficult to link that household's television usage to its digital media usage. This presents a challenge for today's media companies: they want to understand how people are using traditional media such as television in conjunction with the emerging digital media platforms, but they lack detailed data on television viewing behavior. The only readily available data on television viewing on a broad scale is aggregate ratings data.[1] Thus, many companies' data on media consumption are of mixed structure: panel data for digital platforms (on which individual usage is tracked daily) along with aggregate daily or hourly viewership for television and other traditional media. The goal of this research is to address the methodological problem of understanding consumer behavior given this general data structure.

In particular, we develop a Bayesian data-fusion approach to provide inferences for mixed-level (individual-level and aggregate) data and apply it to a case study on multiplatform media consumption. We begin by specifying a model for individual behavior. In specifying the individual-level model, our primary interest is to understand the correlations in usage among platforms so that we can identify whether usage of two (or more) platforms is positively or negatively related. This motivates our use of a hierarchical multivariate logit model featuring a vector of platform-specific intercepts for each consumer. Within this structure, we allow for negative or positive correlations among the platform intercepts for each user as well as negative or positive correlations among daily error terms. This enables us to distinguish the daily substitution among platforms (i.e., use of one platform on a given day may be correlated with use of another platform on that same day) versus the long-term positive correlations that researchers typically find in media consumption behavior (i.e., people who consume more often on one platform may be more likely to consume more on other platforms over the long run). We find it somewhat amusing, therefore, that this is one of those rare occasions in which elements of covariance matrices are truly parameters of interest to understand the relationship between platforms (and not just nuisance parameters to soak up unexplained variation). With these parameters that measure the relationships among media channels, content providers can gauge whether new platforms are detracting from or enhancing consumption on existing platforms, and this information is critical for determining whether to invest in a new channel. In our case study, we demonstrate this by using the estimated model to forecast the impact on viewership by changing which content is available on the mobile channel, the most recent platform entrant.

If data on individual-level media consumption matched across all platforms were available, we could straightforwardly estimate this hierarchical model using Markov chain Monte Carlo (MCMC) methods. However, when one (or more) platforms is only measured in aggregate, model estimation is not straightforward. Our goal is to relate the aggregate behavior observed for one channel to the observed individual-level behavior for the other channels, thus enabling us to make a coherent inference about the joint distribution of behavior across all channels. This is akin, but not identical, to the "direct approach to data fusion" that Gilula, McCulloch, and Rossi (2006, p. 74) propose: They define data fusion as "the problem of how to make inferences about the joint distribution of two sets of random variables (hereinafter called the 'target' variables) when only information on the marginal distribution of each set is available." In their setting, they observe the target variables for two separate sets of people. As they point out, the joint distribution is not identified without additional data and model assumptions. To identify the joint distribution, they use a set of common variables that are observed at the individual level in both data sets under the assumption that the variables to be linked are independent conditional on the common variables.

Our setting varies substantially from that of Gilula, McCulloch, and Rossi (2006). We do not have a set of individual-level linking variables that span the platforms. In addition, for one of the platforms, we do not observe any data at the individual level. In this way, our (dis)aggregate data problem combines elements of the data-fusion problem with the problem of estimating an individual-level model from aggregate data. To handle the aggregate data, we compute the likelihood for it by integrating over the individual-level likelihood of our core model to obtain the marginal likelihood. We accomplished this within the MCMC framework through data augmentation (Tanner and Wong 1987), in which, within the Gibbs sampler, we simulate (many) sets of individual-level behaviors that are exactly consistent with the observed aggregate data. This approach is similar to methods in the marketing literature for estimating individual-level choice models from aggregate data (Albuquerque and Bronnenberg 2009; Chen and Yang 2007; Musalem, Bradlow, and Raju 2008), albeit in a context of a multivariate model rather than a choice model and in a situation in which there is mixed aggregate and disaggregate data.

[1]Some television ratings providers, such as Nielsen, Kantar, and GfK, do collect and sell both television viewing and Internet usage data for a limited number of households; they collect such data by recruiting users to join a media panel and providing them with devices to measure their media consumption across platforms. This is typically referred to as "user-centric" measurement. Such panels are not only expensive to run but also subject to significant selection bias and panelist compliance problems. Here, we focus on lower-cost, more reliable "site-centric" data (as in Zheng, Fader, and Padmanabhan 2012).

To link the aggregate and disaggregate data together, we exploit the longitudinal nature of our data (a feature not present in the problems that Gilula, McCulloch, and Rossi 2006 discuss). Specifically, we exploit the repeated measures of consumer behavior across platforms to identify the intraday correlations between platforms. Note that we cannot identify the correlations between platform-specific intercepts for the aggregate channel and each of the other channels because we do not have any linking variables. However, we demonstrate, through a parameter recovery study using synthetic data, that the intraday correlations between platforms can be recovered even when usage for one of the platforms is only measured in aggregate. Furthermore, our approach enables us to make inferences on the basis of all the observed data and to fully characterize the posterior uncertainty that arises even though we do not have the ideal individual-level data, as is the case in other Bayesian approaches to missing data (Little and Rubin 2002) but not in more ad hoc approaches to data fusion.

Although we motivate the development of this model from the perspective of media consumption, we emphasize that the general data structure we describe here occurs in many other settings, and thus researchers could apply these statistical methods in other contexts. It is remarkably common for marketers to observe consumer behavior across multiple platforms, and they often observe these data in disaggregate for some platforms and in aggregate for others. For example, retailers traditionally have not tracked individual customers' visits or purchases at physical stores but do have accurate data on aggregate traffic and sales for each store. By contrast, retailers do regularly track visits and purchases at the individual level in online stores. The same model we propose for multiplatform media consumption could be applied directly to this type of multiplatform online/offline store visit behavior or purchase behavior. Surprisingly, despite the ubiquitous nature of this data structure in marketing, no direct method exists to perform the mixed aggregate and disaggregate data fusion that we present here.

APPLICATION TO WORLD CUP MEDIA CONSUMPTION

To preview the interesting nature of our applied case study and to motivate the data structure for which our mathematical model is built, we briefly describe the data here (and offer more details when we present the case study). We demonstrate our method using data on multiplatform media consumption during the 2010 FIFA World Cup. The core data in our case study are observations of digital media usage for a random sample of 2,000 ESPN users based primarily in North America. For each of these users, we observe daily consumption across three platforms: a traditional website (ESPN.com), online streaming video, and a mobile website.[2] We augment this with aggregate ratings for World Cup games broadcast on television during the same period. Although our data and case study are limited to a particular type of content at a particular point in time, most major content providers collect data that are similar to the data we use, and other content providers could use the basic modeling approach we propose to measure and understand the changing relationships among platforms as media consumption behavior evolves and new platforms are introduced.

RELATED LITERATURE

Media planners have shown interest in how users interact with multiple platforms, particularly as new media platforms have begun to proliferate (see, e.g., Franz 2000). A key question in the literature is whether consumers use multiple media platforms at the same time.[3] Prior theoretical work has shown that rational media consumers use multiple platforms under certain circumstances (Parker and Van Alstyne 2005). However, most of the empirical work in this area only *suggests* that individual users might be using multiple media platforms. For example, Joo, Wilbur, and Zhu (2012) find that there is a relationship between the airing of television advertisements and aggregate online search behavior, suggesting that individual users must be using television and search at the same time. Similarly, aggregate marketing mix models have shown synergy effects between advertising channels (Naik and Peters 2009; Naik and Raman 2003), and the most likely explanation for this synergy is that individual consumers are viewing content on multiple channels and that the cross-channel repetition is highly effective. Thus, the issue of whether viewers use multiple media platforms is of great interest to media planners, but few empirical studies have investigated individual-level multichannel media consumption directly, partly because of past data scarcity. Although it is not our primary focus, this is one of the first academic articles to examine multiplatform media consumption using directly observed behavior. Prior survey-based studies of multiplatform user behavior have also focused on advertising exposures and advertisers' questions (e.g., "Which advertisements should I place and where should I place them to maximize sales?"), whereas our case study focuses on the media planning problem from the content provider's perspective (e.g., "On which platforms should I place content?").

Although our study does contribute to the body of work on multichannel media consumption, our focus is more methodological, in the vein of empirical work that has developed models for closely related data structures. Our core data structure is one in which consumers engage in multiple activities over time, and thus, our individual-level model is structurally similar to those that examine which websites users visit over time (Danaher 2007), which categories consumers purchase from over time (e.g., Manchanda, Ansari, and Gupta 1999), and which distribution channels a customer uses over time (e.g., Ansari, Mela, and Neslin 2008). However, although all these studies model a person's multiplatform usage over time, none has tackled the issue of when data on usage of one (or more) of the platforms is only

[2]Indeed, our data are, in a sense, "as good as it gets" for media content providers because many ESPN users self-identify (i.e., log in) as they use content across the three platforms we observe. However, as we discuss in the conclusion, the approach we propose could be extended to a situation in which individual-level data are unmatched, that is, one in which there is a distinct group of users for one or more platforms.

[3]The vocabulary to describe this behavior is still evolving. The theoretical literature has used "multihoming" to refer to a scenario in which consumers use one or more platforms but not necessarily at the same time (Parker and Van Alstyne 2005), whereas other work has proposed "multiplexing" to refer to consumers who use two media platforms simultaneously (Lin, Venkataraman, and Jap 2011).

available in aggregate. This is somewhat surprising given how frequently this general data structure occurs in practice.

As we mentioned previously, prior research has addressed the problem of estimating multinomial choice models from aggregate data, most notably in the generalized method of moments framework using the best linear prediction method (Berry, Levinsohn, and Pakes 1995; Nevo 2000). Alburquerque and Bronnenberg (2009) extend this approach to the situation in which both aggregate shares over time and aggregate data on purchase set size and penetration rates are available. These articles all propose moment conditions that researchers can use to estimate an individual-level multinomial choice model from different types of aggregate data. By contrast, we use Bayesian data augmentation to relate the aggregated data to the marginal likelihood of the individual-level model. In this way, we are more methodologically akin to Chen and Yang (2007), who estimate a multinomial logit model using Bayesian data augmentation, and to Musalem, Bradlow, and Raju (2008), who show that data augmentation can be used not only to estimate an individual-level model from aggregate choice data but also to handle the situation in which covariates are only observed in aggregate (e.g., estimating which consumers have access to a coupon when only aggregate data on how many coupons were distributed is available). However, none of these prior approaches to estimating individual-level models from aggregate data address how to estimate a multivariate outcome (rather than a choice) or how to proceed when some outcomes are observed at the individual level and others are only observed in aggregate.

We have structured the remainder of the article as follows: In the next section, we describe a model that enables us to fuse the aforementioned aggregate and disaggregate data. We then report a synthetic data study that explores the extent to which correlations can be recovered when one platform is measured in aggregate. To illustrate the method, we report the application to the FIFA World Cup data, including a set of forecasts predicting ESPN usage had the company not provided coverage (either fully or in part) on its new mobile platform. We conclude with a discussion of potential applications and extensions of the method.

A HIERARCHICAL BAYESIAN MULTIVARIATE LOGIT MODEL FOR MULTIPLATFORM USAGE

Model for Disaggregate Multiplatform Behavior

We focus first on developing a model for disaggregate data in which, for each user in the sample, we observe a vector of binary outcomes, y_{ikt}, indicating whether a user, $i = 1, ..., N$, accessed content on a given platform $k = 1, ..., K$ on day $t = 1, ..., T$. After laying out the model for disaggregate data, we discuss modifications for the situation in which one or more platforms are observed only in aggregate.

We model y_{ikt} with a multivariate hierarchical logistic regression:

$$y_{ikt} \sim \text{Bernoulli}(p_{ikt}) \tag{1}$$

$$\text{logit}(p_{ikt}) = \mu_{ik} + x_{kt}\beta_k + e_{kt}, \tag{2}$$

where x_{kt} is a vector of covariates describing the content available on platform k on day t, which is multiplied by a vector of platform-specific coefficients β_k. The residual appeal of the platform after controlling for the covariates, μ_{ik}, is user-specific, so this can thus be considered a mixed-effects model.

The vector $\mu_i = (\mu_{i1}, ..., \mu_{iK})$ is assumed to be normally distributed across the population:

$$\mu_i|\mu, \Sigma_\mu \sim N_K(\mu, \Sigma_\mu), \text{ i.i.d.} \tag{3}$$

We include the parameter μ_i to accommodate differences in overall usage rates among active users across the platforms. The matrix Σ_μ captures the covariance among consumers' propensities to use each of the platforms and, as we mentioned previously, is one of the sets of model parameters of greatest interest because it represents the covariance among baseline usage propensities across platforms over time.

The error term $e_t = (e_{1t}, ..., e_{Kt})$ from Equation 2 is also modeled as a multivariate normal:

$$e_t|\Sigma_e \sim N_K(0, \Sigma_e), \text{ i.i.d.}, \tag{4}$$

which allows for correlations among the propensities to use each of the platforms on a given day through Σ_e.[4]

This defines our core model for multiplatform consumer behavior. However, in many marketing data sets, we observe a large number of users who are completely inactive during the observation period. If we include these inactive consumers in the estimation of Σ_μ, we would likely find strong positive correlations between platforms induced by this one group of consumers who are inactive on all platforms. To avoid this, we model each user as either completely inactive (zero on all platforms for each and every day) or active with probability p_{active}, a classic "spike-at-zero" mixture model (Morrison and Schmittlein 1988). We use I_i to indicate (latently) whether user i is active. This aspect of the model would not be necessary in situations in which most of the users have some observed activity.

If $y_{it} = (y_{i1t}, ..., y_{iKt})$ is the vector of observed platform usage for user i in period t, the full likelihood for the hierarchical model is given by

$$\begin{aligned} \ell(\{y_{it}|I_i = 1\}) &= \prod_i \left[\prod_t (y_{it}|\mu_i, \beta_k, e_t)(e_t|\Sigma_e)\right](\mu_i|\mu, \Sigma_\mu) \\ &= \prod_i \left\{\prod_t \left[\prod_k (p_{ikt})^{y_{ikt}}(1 - p_{ikt})^{1-y_{ikt}}\right] \times \right. \\ &\qquad \left. N_K(e_t|0, \Sigma_e)\right\} N_K(\mu_i|\mu, \Sigma_\mu). \end{aligned} \tag{5}$$

In summary, the model we propose for consumers' disaggregate media consumption across multiple platforms is a hierarchical multivariate logit model with a spike at zero for a consumer's use of multiple media platforms. Our model extends the basic multivariate logit framework (Glonek and McCullagh 1995) by decomposing individual-level cross-day platform effects through Σ_μ from within-day cross-platform effects through Σ_e. This structure enables us to determine the short- and long-term correlations between media platforms, which, as we explained previously, is critical for media

[4]This specification assumes that e_t is independent of e_{t+1}, although this raises the question whether this assumption is warranted. That is, is viewership likely to be higher today conditional on viewership being high yesterday? To confirm this assumption in our case study, we calculated the lag-one autocorrelation of the residuals for each platform. None of the residual sequences showed significant autocorrelation.

planners who want to understand how consumers combine the use of channels. The hierarchical Bayesian framework also enables us to determine which individual users within the sample are most likely to watch each platform through the estimates of μ_{ik}. We tested a similar hierarchical multivariate probit specification (Chib and Greenberg 1998; Rossi, Allenby, and McCulloch 2005) and found it to have similar fit.

Model for Mixed Aggregate and Disaggregate Data

As we described previously, companies often do not observe individual-level data, y_{ikt}, for the full set of platforms of interest. Typically, they have detailed panel data suitable for estimating the aforementioned model for newer digital platforms (e.g., website visits in the case of media consumption, online visits in the case of a multichannel retailer), but data on usage of traditional platforms (e.g., television, brick-and-mortar store purchases) are only available in aggregate.

Although our method can be applied when more than one channel is observed in aggregate, for simplicity of exposition, we assume that there is only one platform observed in aggregate and that this is the Kth platform. For this platform, we do not observe y_{iKt} but instead observe $Y_{Kt} = \Sigma_i y_{iKt}$ for each time period. We can obtain the likelihood for the observed disaggregate and aggregate data by integrating the likelihood of the model over all possible values of $\{y_{iKt}\}$ that meet this constraint as follows:

$$(6) \qquad \ell\left[Y_{Kt}, (y_{it}|I_i = 1)\right] =$$

$$\int_{\{y_{iKt}\} s.t Y_{Kt} = \sum_i y_{iKt}} \left\{\prod_i \left[\prod_t (y_{it}|\mu_i, \beta_k, e_t)(e_t|\Sigma_e)\right](\mu_i|\mu, \Sigma_\mu)\right\},$$

where the integral is taken over all possible values of the set $\{y_{iKt}\}$ that meet the sum constraint implied by the observed aggregate viewership. This is the likelihood we use in making posterior inferences.

We can estimate this model in the Bayesian framework through data augmentation of y_{iKt} using standard MCMC methods (Tanner and Wong 1987). Instead of treating y_{iKt} as a nuisance parameter, the data augmentation approach treats "missing" data such as y_{iKt} as a parameter to be estimated conditional on the constraint. Under the MCMC framework, this results in draws from the posterior of all the model parameters as well as from the posterior of y_{iKt} conditional on the model structure and all observed data.

We note that when the Kth platform is only observed in aggregate, the last row and column of Σ_μ, corresponding to the covariance between the propensity to use the aggregate platform and the other platforms on a given day, is not identified. Intuitively, we never observe which individual consumers are using the Kth platform on a given day, so it is impossible to estimate the covariance in the platform intercepts between the aggregated platform and the others. However, as we demonstrate in a simulation study reported in the next section, the covariance in Σ_e is identified through the repeated measures over time. Consequently, we fix the covariance elements of Σ_μ associated with the Kth platform to zero. Note that this model still allows for an i.i.d. Bernoulli error for daily television usage through Equation 1.

We obtain posterior samples for the population-level parameters (μ, β_k, Σ_μ, and Σ_e) and the individual-level parameters (μ_i as well as the "missing" sets of individual-level viewership $\{y_{iKt}\}$) using an MCMC sampler implemented in WinBUGS (Spiegelhalter, Thomas, and Best 1999). We use diffuse but proper priors, as we describe in the Web Appendix (www.marketingpower.com/jmr_webappendix). The Web Appendix also includes details on how to specify the margin constraint in WinBUGS. All code is available from the authors on request.

In the next section, we report the results of a parameter recovery study showing that the approach makes accurate inferences about the true population-level parameters, particularly the correlations Σ_μ and Σ_e, when the data are generated according to the true model. We also explore parameter recovery when the model is slightly misspecified in an important way that enables us to test a fundamental assumption that we made.

PARAMETER RECOVERY

The first parameter recovery study tests the empirical identification of the individual-level model when one platform is observed in aggregate. For this study, we generated a relatively modestly sized data set (N = 200, K = 3, J = 30) according to the true individual-level model and then aggregated the observed individual-level y_{ikt} for platform K = 3 to Y_{3t}. As Table 1 illustrates, the parameters of the population-level distributions, particularly the correlations in Σ_μ and Σ_e, can all be reasonably recovered (i.e., the posterior covers the values of the population-level parameters used to generate the synthetic data). Furthermore, the posterior is significantly narrower than the associated priors, indicating that the likelihood contributes to the posterior; that is, there is some "learning" from the data. This suggests that the parameters we estimate, particularly the correlations in people's long-term propensity to use the various platforms in Σ_e, are identified by the data/likelihood combination as well as by the other model parameters.

However, as we discussed in the previous subsection, the correlations of the individual-level platform intercepts between the aggregated platform K and the other platforms are not identified; therefore, without any other identifying information, we fix the correlations to be zero. Yet what if this is a misspecification and the platform intercepts for the Kth platform are correlated with the intercepts for one of the disaggregated platforms? Does this negatively affect (bias) parameter recovery for the identified parameters? To assess this, we generated a second synthetic data set (N = 500, K = 3, J = 30) in which the correlation in Σ_μ between the first and the Kth platform is nonzero and set equal to .4, whereas all other correlations were set to zero. As Table 2 shows, we correctly recovered all the parameters (except the unidentified correlation that we did not estimate), and our ability to recover the empirically identified parameters was not degraded. Thus, we find evidence that there is minimal bias in the parameter estimates even when there is a correlation between the platform intercepts (which, as we stated previously, the data did not identify).

APPLICATION TO FIFA WORLD CUP VIEWING BEHAVIOR

To demonstrate the model and its use for forecasting and understanding media channel interplay, we applied it to a

Table 1
ESTIMATED MODEL PARAMETERS FOR SYNTHETIC DATA GENERATED ACCORDING TO THE TRUE MODEL

Parameters	*True*	*Posterior Mean*	*2.5th Percentile*	*97.5th Percentile*
$\mu[1]$	–.500	–.696	–.978	–.435
$\mu[2]$	–.300	–.120	–.379	.137
$\mu[3]$	–.800	–.695	–.999	–.393
$\Sigma_\mu[1, 1]$	.500	.458	.327	.630
$\Sigma_\mu[1, 2]$	.200	.163	.059	.282
$\Sigma_\mu[2, 2]$	.500	.605	.390	.871
$\Sigma_e[1, 1]$	.700	.535	.306	.917
$\Sigma_e[1, 2]$	.350	.215	.041	.462
$\Sigma_e[1, 3]$	.280	.251	.030	.559
$\Sigma_e[2, 2]$	.700	.433	.230	.772
$\Sigma_e[2, 3]$	.210	.196	–.010	.478
$\Sigma_e[3, 3]$	.700	.759	.440	1.285

Table 2
ESTIMATED MODEL PARAMETERS FOR SYNTHETIC DATA GENERATED FROM A MISSPECIFIED MODEL

Parameters	*True*	*Posterior Mean*	*2.5th Percentile*	*97.5th Percentile*
$\mu[1]$	–.500	–.277	–.525	.037
$\mu[2]$	–.300	–.165	–.463	.093
$\mu[3]$	–.800	–.833	–.966	–.645
$\Sigma_\mu[1, 1]$	.500	.482	.403	.572
$\Sigma_\mu[1, 2]$	.000	.019	–.046	.083
$\Sigma_\mu[1, 3]$	.200	N.A.	N.A.	N.A.
$\Sigma_\mu[2, 2]$	.500	.633	.530	.744
$\Sigma_\mu[2, 3]$	.000	N.A.	N.A.	N.A.
$\Sigma_\mu[3, 3]$	.500	N.A.	N.A.	N.A.
$\Sigma_e[1, 1]$	.700	.652	.408	1.025
$\Sigma_e[1, 2]$	.000	.026	–.214	.267
$\Sigma_e[1, 3]$	.000	.070	–.062	.225
$\Sigma_e[2, 2]$	.700	.727	.449	1.159
$\Sigma_e[2, 3]$	.000	.048	–.092	.208
$\Sigma_e[3, 3]$	.700	.262	.066	.425

Notes: N.A. = not applicable.

data set on the multiplatform media consumption of 2,000 ESPN registered users[5] for each day from June 4 to July 11, 2010, during the 2010 FIFA World Cup, the most watched sporting event in the world. The observation window includes the week before the start of the World Cup tournament, which began on June 11, 2010, and concludes with the final championship game on July 11, 2010. This project was part of a larger initiative (with many participating media research firms) called ESPN XP (Crupi 2010), designed to help ESPN and their advertisers better understand cross-platform user behavior.

Because the focus of this ESPN project was to shed light on multiplatform behavior, and because the majority of U.S. users did not yet have smartphones or other mobile devices at the time of the tournament, the target population from which we sampled were users who were observed to use mobile services (for any ESPN content) sometime in the year before the tournament. These users are important to the company because they represent what we might consider the vanguard of mobile users. For each user on each day, we observe a binary vector indicating whether he or she watched or read soccer-related content on each of three media platforms: (1) ordinary digital content on the regular magazine-format website (ESPN.com), (2) live or archived streaming video of full games available on the ESPN3.com platform (a separate website from ESPN.com),[6] and (3) ESPN Mobile (a site specifically formatted for mobile phones and tablet devices with a smaller screen that includes mobile friendly features such as "ESPN Gamecast"). The sample is not conditional on the user being a soccer fan or even having consumed soccer content before the tournament; thus, it is not surprising that approximately half of the sampled users, who are primarily based in the United States, did not access any soccer-related content during the tournament (therefore the need for the aforementioned spike at zero). Figure 1 shows the total daily reach for each platform.

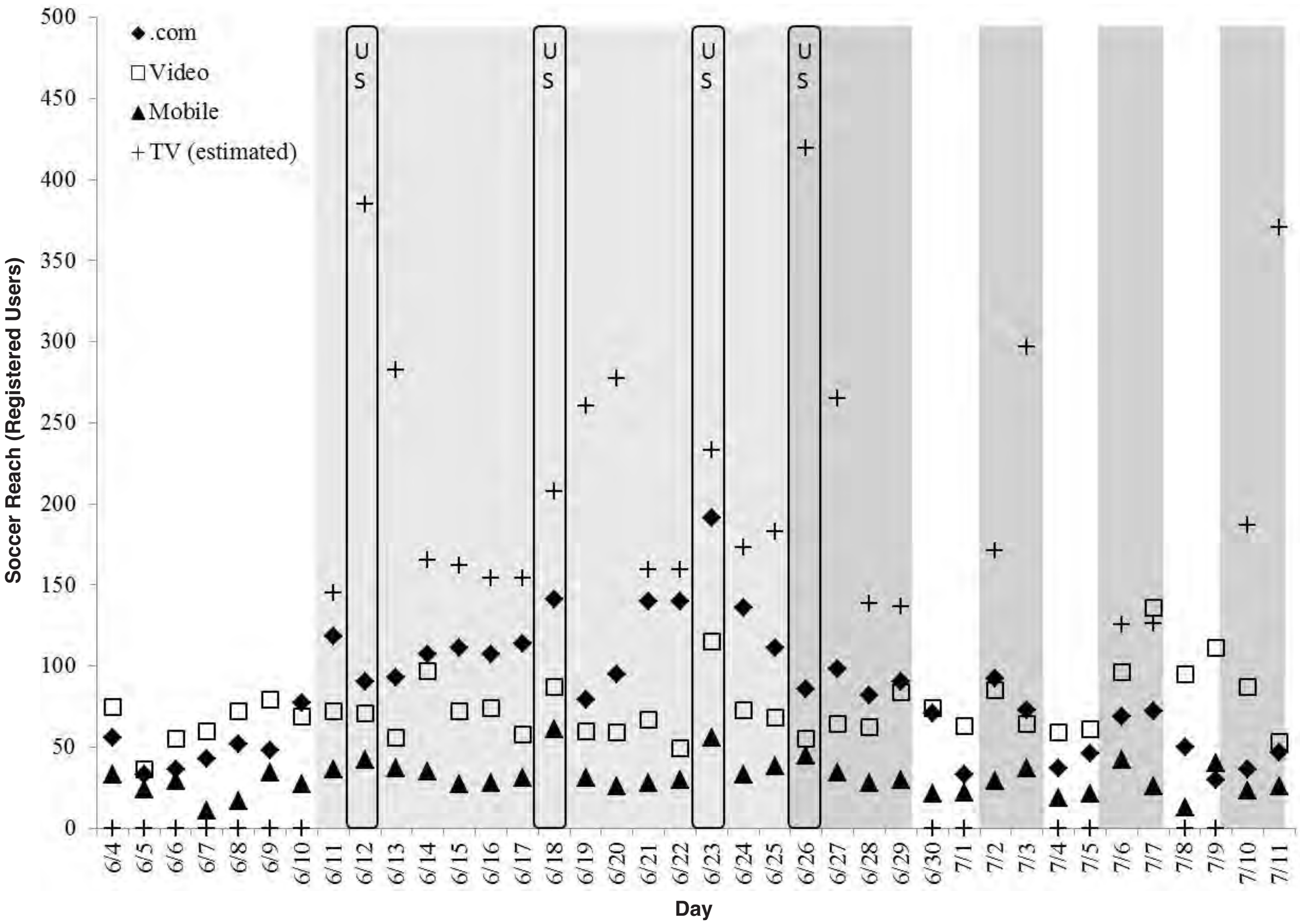
We combine these digital media data with aggregate data on television ratings provided by The Nielsen Company. Specifically, Nielsen provided its estimate of the total fraction of U.S. households that watched any of the televised English-language broadcasts of World Cup games for each day during the tournament.[7] As we discussed previously, our model is designed for the situation in which we observe Y_{Kt}, the total number of people who watched television among our panel of digital users. To accommodate the ratings data within our framework, we make one assumption: that the fraction of our digital users who watched the television broadcast of the World Cup approximately matches the fraction of total U.S. households who watched on television as reported by Nielsen. That is, we assume our sample of digital platform users who watched the World Cup on television to be the same proportion as the general population. Because television is currently the most popular media platform, we find this assumption reasonable. Note, however, that Y_{Kt} is directly observed in many other applications. For example, some media companies may have access to television consumption data specifically for digital media users; or, in another context (to which our methodology applies), a multiplatform retailer may know the total sales in brick-

[5]Some might question whether registered users are the appropriate group on which to focus. We chose this group for two reasons. First, they can be tracked longitudinally over time and across platforms (because they log in), which is critical to the study of multiplatform behavior. Second, they are fairly representative of ESPN users as a whole because ESPN has designed their platforms to encourage a majority of their users to log in (e.g., for fantasy sports games). The only other practical alternative for studying multiplatform media behavior is to analyze data from an opt-in, multiplatform media consumption panel (e.g., from Nielsen), which raises another set of selection issues.

[6]Although both ESPN.com and ESPN3.com are viewed on the Internet, from the point of view of ESPN and its users, ESPN3.com is a distinct platform with an entirely different interface than ESPN.com and is only available to users who subscribe to certain cable providers. Unfortunately, from a back-end technology perspective, the same server hosted video for both sites, and ESPN commingled data on usage of ESPN3.com and streaming video embedded in ESPN.com. However, because nearly all the video on ESPN.com was in the form of short clips, with longer full-game videos reserved for ESPN3.com, we were able to approximate ESPN3.com usage by only counting the user as having watched ESPN3.com if he or she viewed a video for more than three minutes.

[7]We focused our analysis specifically on English speakers because the ESPN.com, ESPN3.com, and ESPN Mobile audiences are English-language oriented. We excluded from our analysis broadcasts on Spanish-language television and the relatively fewer users of ESPN's Spanish-language website, ESPN Deportes.

Figure 1
DAILY SOCCER REACH FOR ESPN MEDIA PLATFORMS DURING THE 2010 WORLD CUP



Notes: For 2,000 registered users with mobile usage in the past 12 months. Days on which tournament games occur are shaded. Days on which the U.S. team was playing are highlighted in boxes.

and-mortar stores precisely. When applying the method, we suggest analysts carefully consider this assumption in their applications.

Covariates to Control for Tournament Content

We selected the covariates we specify in x_{kt} to account for the relationship between the tournament content and viewing behavior. Without controlling for spikes in reach that are driven by tournament content (as Figure 1 illustrates), it would be difficult to interpret the correlations between platforms that are our primary interest. Although our covariates are specific to tournament content, in other applications, a similar set of covariates could be included to represent seasonality, promotions, and other outside factors that are known to influence the behavior being studied.

It is commonly understood in practice that television reach is greater on weekends and online reach is greater on weekdays, reflecting the intuition that television is generally more accessible on weekends and that viewers turn to online coverage when they are at work. Thus, we included a dummy variable indicating whether the day was a weekend, which enables us to confirm previously observed weekend/weekday effects for television and online as well as to determine whether the new mobile platform displays a similar pattern to online or television.

Regarding the tournament itself, there is little prior literature addressing what might drive viewership in this setting. We assume that the teams that play on a given day have a substantial impact on reach; therefore, we introduce a set of covariates to control for such effects, including (1) a dummy for whether the U.S. team played, (2) a dummy for whether any of the three top teams (Brazil, Spain, and the Netherlands, according to the May 2010 FIFA soccer rankings) played, and (3) a dummy indicating whether any of the three teams with strong cultural connections to the United States (England, Mexico, and Australia) played.[8] Assuming

[8]We also explored the possibility of using a shrinkage model to estimate a media attraction effect for every team, but this proved difficult with the available data. A total of 32 teams began the tournament and were assigned into eight groups of 4 teams each and played a round robin with the other teams in their group. From the results of this group stage, half the teams (two from each group) proceeded to a knockout-style, one-and-done elimination tournament. This results in a tournament schedule in which half the teams only play on three days, sometimes perfectly overlapping with another team, giving us limited information from which to infer attractiveness effects for each of them.

that a greater number of games and games that are more critical may drive additional reach, we also include variables for the total number of games on a given day and the number of teams that would be eliminated from the tournament on that day if they failed to win—that is, the number of teams that must "win or go home."

Although we present and interpret the parameter estimates for these effects, we caution that the estimates for models of this type are subject to the usual potential biases due to collinearity, missing variables, missing interactions, and other misspecifications (although we did guard against them as much as possible). In general, this type of regression is commonly used in practice and, like any other regression analysis with nonexperimental data, we find that the parameter estimates for these effects provide some actionable insights for a media company that aims to understand the drivers of multichannel media viewership. For example, as we discuss when we report the parameter estimates, understanding whether mobile usage is higher or lower on weekends (controlling for other aspects of the content) can provide insight into whether weekend mobile coverage should be emphasized or deemphasized.

As in the synthetic data parameter recovery studies, we estimated the model in WinBUGS. Posterior inferences are based on 50,000 draws from the posterior. We discarded the first 50,000 draws on the basis of trace plots and Gelman–Rubin diagnostics against a second chain run from an independent starting point (Gelman and Rubin 1992). Furthermore, as is appropriate for all complex models, we performed two types of model assessment, in sample and out of sample. For our out-of-sample assessment, we performed two analyses, one temporal (omitting the final days) and the other cross-sectional (using one set of viewers to predict another), thus running a stringent set of model tests showing that this model is appropriate for our media-consumption data.

Model Assessment

Our assessment of model fit focuses on a series of posterior predictive checks at varying levels of disaggregation (Gelman, Meng, and Stern 1996) that evaluate the model's ability to fit features of the data that are important to media planners. Following the usual procedure for computing posterior predictions, we generated 100 posterior predictive data sets using 100 sets of parameters randomly sampled from the posterior draws obtained from the MCMC sampler. (For more detail, see the Web Appendix at www.marketingpower.com/jmr_webappendix.) We then compared the posterior distributions for these statistics with actuals computed from the data and report the quantile of the observed value within the posterior predictive distribution to assess the ability of the model to correctly recover these key statistics.

Fit of multiplatform usage patterns. By comparing the posterior predictions for the percentage of users who use each combination of platforms, we can assess whether the model is picking up the appropriate covariation—that is, cross-platform usage, our central question of interest. The last two columns of Table 3 report the ability of the model to predict the 2^3 contingency table of aggregate usage over the course of the tournament for each combination of the digital platforms. (Because we do not observe television usage at the individual level and thus cannot compute television's actual cross-platform contingencies, we do not include it in this posterior predictive check.) As the data show, the fit is good, with the true values all falling within the .15 and .75 quantiles of the estimated posterior, suggesting that the dual covariance structure adequately captures the covariances we observe among platforms in the ESPN data.

Table 3
COMPARISON OF ESTIMATED AND ACTUAL MULTIPLATFORM USAGE

		Predicted	
	Actual	*Mean*	*Quantile*
No access	.631	.605	.66
Only ESPN.com	.043	.045	.34
Only ESPN3.com	.058	.058	.45
Only ESPN Mobile	.064	.063	.48
ESPN.com and ESPN3.com	.132	.125	.75
ESPN.com and ESPN Mobile	.018	.021	.27
ESPN3.com and ESPN Mobile	.017	.021	.17
All three platforms	.057	.062	.27

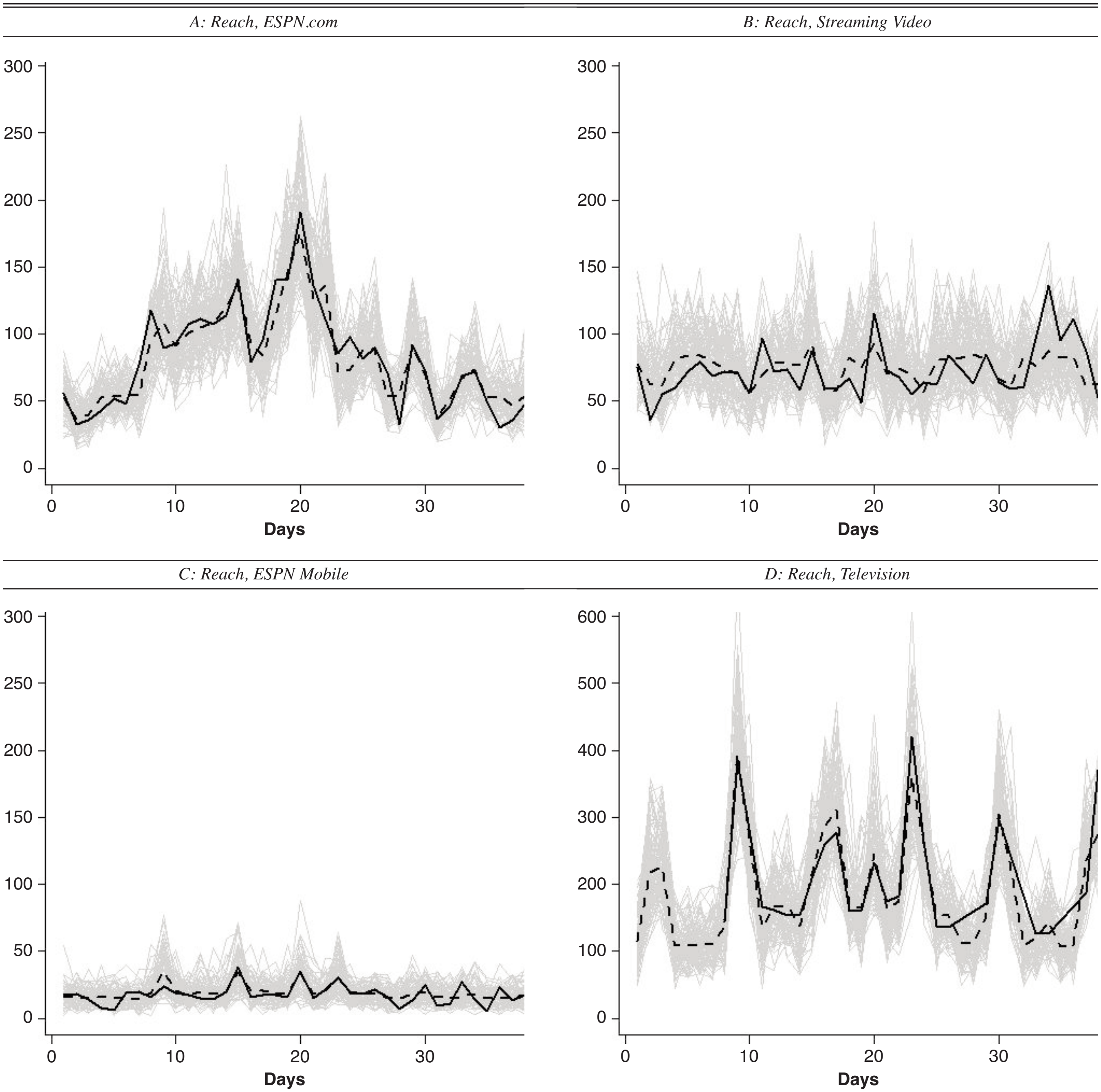
Tracking plots for daily reach. Figure 2 plots the posterior mean prediction for daily reach (solid line) compared with the actual daily reach (dotted line) for all four platforms. We also show posterior uncertainty by plotting the daily reach predictions for the 100 posterior draws with gray lines. These tracking plots show an excellent fit between the model and the data: the overall mean absolute error between the predicted daily reach (posterior mean) and the actual reach is low: .78% for ESPN.com, .71% ESPN3.com, and 1.82% for ESPN Mobile, suggesting that the model adequately captures the major features of the total daily reach in the data, including the day-to-day variation.

In Figure 3, we show that the model also does an excellent job at predicting cumulative reach for the digital platforms. Media planners frequently use cumulative reach (defined as the total number of unique users who have viewed content on a particular platform through to a specific day) to understand how many viewers could be reached over the course of an entire media event such as the World Cup. Note that, as before, we cannot compute the actual cumulative reach for television, because we do not observe which users are viewing on each day; however, we report the model prediction for completeness. The model suggests that cumulative reach for television levels out approximately halfway through the tournament, indicating that most of those who watch a game on television will watch their first game fairly early in the tournament; relatively few viewers wait until the knockout stage to first watch a game.

Figures 2 and 3 demonstrate that the model, estimated from the full 38 days of data, can fit the observed patterns in the data well. In practice, researchers might imagine using the model to make predictions for the future; therefore, to test the model's forecasting accuracy, we reestimated the model using the first 36 days of data and then predicted the response for the final two days (the third place and final games). This corresponds with the scenario in which we estimate the model just after the 36th day of the tournament and then use it to predict reach for the final two days on the basis of the teams that are scheduled to play on those days. Figure 4 shows the predicted daily reach for this "two-ahead" prediction with the two out-of-sample forecasts shown with solid circles for the actual reach and open circles for the pre-

Figure 2

TRACKING PLOTS OF DAILY REACH FOR EACH PLATFORM

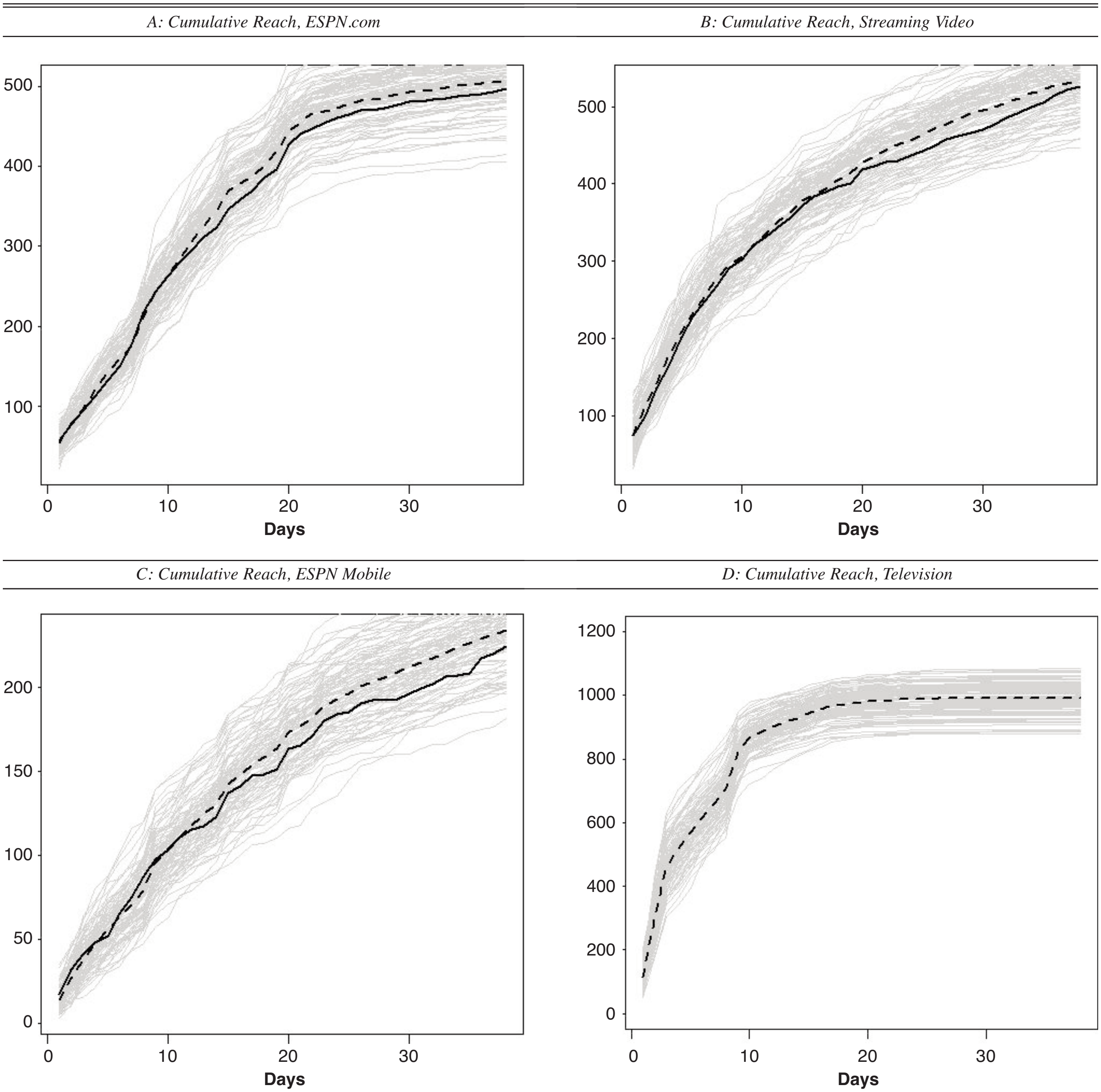


Notes: We draw the mean of the generated statistics with a dashed line and compare it with the actual statistic (computed directly from the sample), drawn with a solid line. To show the forecast uncertainty, we also draw the prediction for each of 100 random draws from the posterior with gray lines.

diction. As the plots illustrate, the predictions are good, and all observed values fall within the prediction error of the model (represented by the gray lines). The model does underpredict television viewing for the finals, although the actual viewership was within the range the model predicted. This underprediction is most likely due to a spike in advertising and consumer interest that is specific to the final game. Because our model incorporates no effects that are specific to the tournament stage, we find this performance quite reasonable.

Although Figure 4 shows a temporal holdout validation, we also wanted to consider a cross-sectional holdout validation. To illustrate how the model fits to a new set of users, Figure 5 shows the tracking plot of daily reach for a different random sample of 2,000 ESPN users we did not use in the estimation. As Figure 5 demonstrates, the fit is reasonably comparable to the fit to the estimation data, suggesting no overfitting for this sample of users. (Average hit rate = .914, .925, and .979 for the ESPN.com, ESPN3.com, and mobile platforms, respectively. We define a person's hit rate

Figure 3
TRACKING PLOTS OF CUMULATIVE REACH FOR EACH PLATFORM



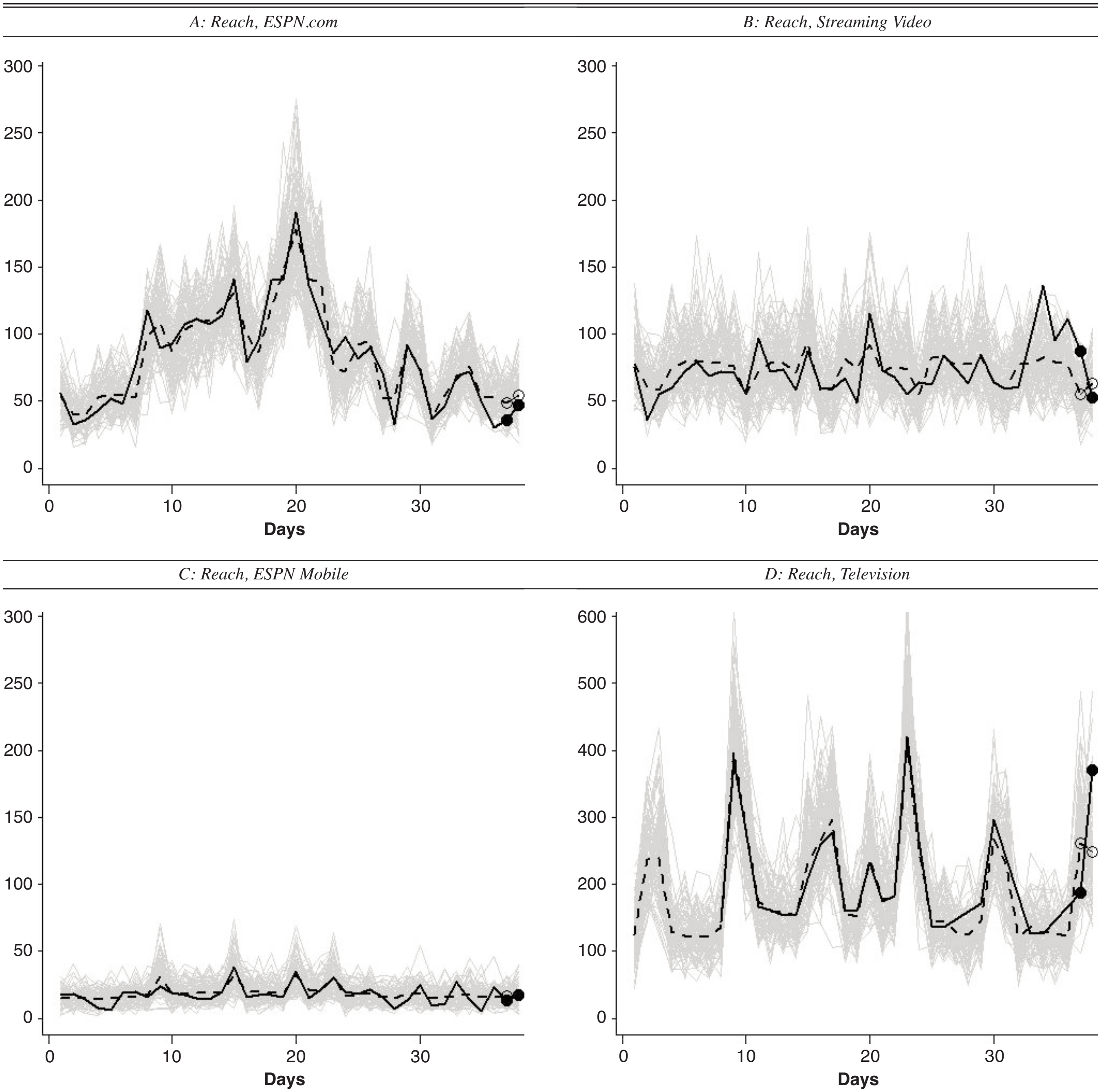
Notes: We draw the mean of the generated statistics with a dashed line and compare it with the actual statistic (computed directly from the sample), drawn with a solid line. To show the forecast uncertainty, we also draw the prediction for each of 100 random draws from the posterior with gray lines. For the television platform only, the aggregate daily usage on game days was observed, and the individual-level usage was not available. Consequently, when we draw the posterior predictive plots for television, we can compute the true daily reach but not the true cumulative reach.

as the proportion of days the model correctly predicts his or her usage incidence.)

Returning to the model estimated from all 38 days with the original 2,000 consumers, our final assessment is intended to determine how well the model fits people's media usage patterns. We compute the hit rate for individuals. As Table 4 indicates, the model does much better than chance at predicting a person's usage with average hit rates of .922, .929, and .990 for all three digital platforms. (We do not report individual hit rates for television because we do not observe individual-level usage on television.) Table 4 also reports the percentage of users who have an average hit rate over the 38 days that is greater than .5 and .95.

In summary, we find that the model we propose fits this multiplatform media usage data well, capturing its key aspects: the aggregate daily and cumulative reach, the pattern

Figure 4
FORECAST OF REACH FOR SEMIFINAL AND FINAL GAME DAYS



Notes: We draw the mean of the generated statistics with a dashed line and compare it with the actual statistic (computed directly from the sample), drawn with a solid line. To show the forecast uncertainty, we also draw the prediction for each of 100 random draws from the posterior with gray lines.

Table 4
HIT RATES FOR INDIVIDUAL USERS

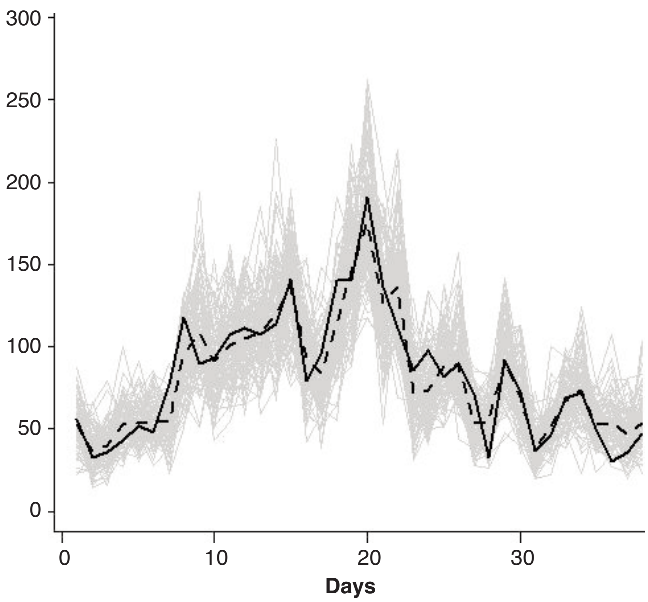
Platform	*Average*	*Percentage of Users with Hit Rate Better Than .5*	*Percentage of Users with Hit Rate Better Than .95*
ESPN.com	.922	96.0	67.7
ESPN3.com	.929	98.0	64.1
ESPN Mobile	.990	99.9	92.8

of co-usage among platforms, and individuals' daily usage. Holdout validations show that the model makes accurate aggregate predictions both for new time periods and for new groups of customers. In the Web Appendix (www.marketingpower.com/jmr_webappendix), we also report several additional posterior predictive checks, demonstrating that the model does a reasonable job recovering the heterogeneity among users (i.e., the number of light and heavy users) and the patterns of usage over time (i.e., how many users "jump in" and "drop out" during the tourna-

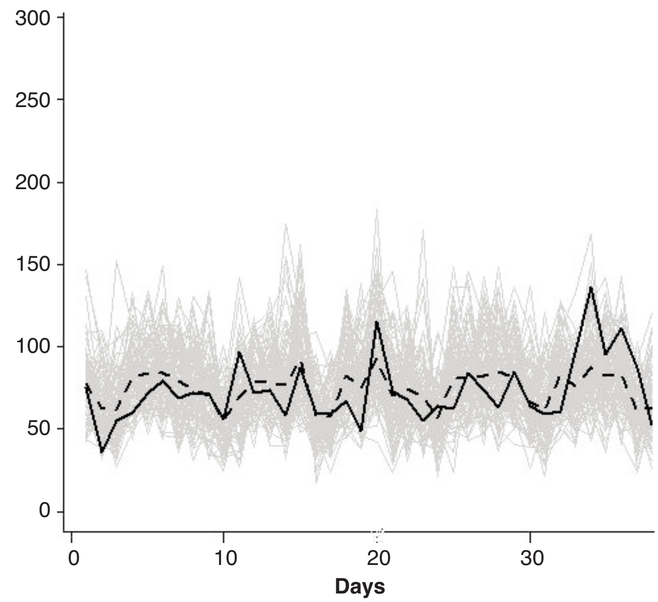
Figure 5

TRACKING PLOTS OF DAILY REACH FOR EACH PLATFORM FOR ANOTHER SET OF CUSTOMERS

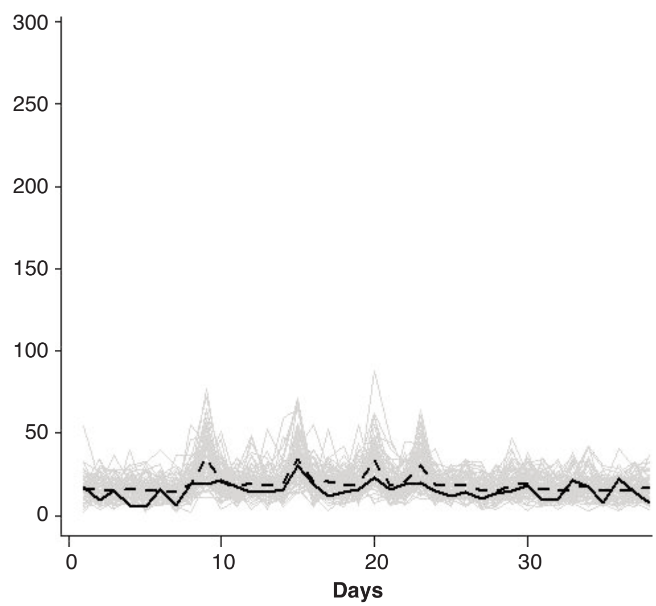
A: Reach, ESPN.com



B: Reach, Streaming Video



C: Reach, ESPN Mobile



Notes: We draw the mean of the generated statistics with a dashed line and compare it with the actual statistic (computed directly from the sample), drawn with a solid line. To show the forecast uncertainty, we also draw the prediction for each of 100 random draws from the posterior with gray lines. The figure on the television platform is the same as the one in Figure 2, because individual-level observations on the television platform are not available.

ment). We suggest that similar model assessments be performed for other applications of the model. With this assurance that the model accurately reflects these key aspects of the data, we next turn to interpreting the estimated parameters of the model.

Parameter Estimates and Implications

To illustrate how the parameters of the model can be interpreted, we discuss, in turn, the parameters that describe the attractiveness of each platform, the heterogeneity around their means, and the correlations of platform preferences over time (Σ_μ) and in intraday platform usage (Σ_e). Following that, we present the platform-specific effects for the tournament characteristics, x_{kt}.

Platform intercepts and user heterogeneity. The posterior mean value for p_{active} is .498, indicating that of our sample, 49.8% of the registered users are "active"; that is, they have some predicted probability of accessing soccer content during the World Cup tournament. Thus, we estimate almost half the sample to be in a spike at zero for each day and each platform, suggesting the necessity for this part of the model. Note that this is consistent with the raw data, in which 61.3% of our sample of 2,000 users has no observed usage on any of the digital platforms. Depending on how the sample is constructed, we would expect this value to vary widely for other data sets, and the spike-at-zero component may not even be needed for data sets in which nearly all users have some observable activity.

The general attractiveness to users of the ESPN.com, ESPN3.com, mobile, and television platforms is reflected by the parameter vector $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$, the population mean of $(\mu_{i1}, \mu_{i2}, \mu_{i3}, \mu_{i4})$. As Table 5 shows, the intercepts for each platform are in the range of –7 to –2 on the logit scale, suggesting that baseline (i.e., on nontournament days) usage of soccer content among active users is still close to zero. This is consistent with the data, in which, for those users who visit at least once, we observe heterogeneous marginal proportions ranging from a low of 2.6% to a high of 100%. Unsurprisingly, we find that television is the most popular platform at baseline ($\mu_4 = -2.06$).

Managerially, it is important to keep in mind that these estimates are based on a sample of users who have been known to use the mobile platform (possibly for non-soccer-related content). We deemed this group to be most relevant because they represent the vanguard of mobile device users who had used the platform prior to 2010. It is notable, then, that they do not strongly prefer the mobile platform over the others. Indeed, the mobile platform has the lowest estimated population mean for the intercept ($\mu_3 = -6.33$), indicating that on days when tournament games are not being played, users are very unlikely to access soccer content on their mobile device. We would expect that this may change over time as consumers become more familiar with mobile devices.

Table 5

ESTIMATED MODEL PARAMETERS: INTERCEPTS AND ERROR STRUCTURE (μ, Σ_μ, Σ_e)

Parameters	*Mean*	*2.5th Percentile*	*97.5th Percentile*
Proportion of Active Users (p_{active})			
p_{active}	**.498**	.455	.536
Population Mean for Platform Intercepts (μ)			
ESPN.com	**–5.01**	–5.01	–4.66
ESPN3.com	**–3.81**	–3.82	–3.49
ESPN Mobile	**–6.33**	–6.31	–5.96
Television	**–2.06**	–2.06	–1.66
Correlation for Platform Intercepts (Σ_μ)			
ESPN.com/ESPN3.com	**.795**	.749	.835
ESPN.com/ESPN Mobile	.056	–.065	.171
ESPN3.com/ESPN Mobile	.028	–.089	.144
Variance for Platform Intercepts (Σ_μ)			
ESPN.com	**6.358**	5.279	7.550
ESPN3.com	**4.233**	3.326	5.023
ESPN Mobile	**5.012**	3.994	6.340
Correlation for Daily Error Terms (Σ_e)			
ESPN.com/ESPN3.com	–.030	–.385	.329
ESPN.com/ESPN Mobile	–.153	–.244	.512
ESPN.com/television	.083	–.321	.466
ESPN3.com/ESPN Mobile	.114	–.275	.476
ESPN3.com/television	–.140	–.508	.269
ESPN Mobile/television	.016	–.404	.436
Variance for Daily Error Terms (Σ_e)			
ESPN.com	**.089**	.051	.149
ESPN3.com	**.105**	.062	.172
ESPN Mobile	**.105**	.053	.191
Television	**.067**	.037	.117

Notes: When the posterior interval does not contain zero, the posterior mean appears in boldface.

We also find, unsurprisingly, substantial variation around these population means, with standard errors for the population distribution in the range of 2.0–2.5, indicating that some users are significantly more (or less) likely to access certain platforms [$\text{diag}(\Sigma_\mu) = (6.358, 4.233, 5.012)$]. In contrast, the variances in Σ_e are small [$\text{diag}(\Sigma_e) = (.089, .105, .105, .067)$], indicating that after we account for the user's general propensity to use a platform, there is little residual error in the daily usage probabilities (other than that driven by the aggregate tournament effects).

Correlation structure. The correlation structure between the channels across time and within each day is one of the central areas of interest to media planners. As we described previously, we summarize such long-term and daily usage effects in two ways. The first is the covariance among users' propensities to use each of the platforms over the course of the tournament, captured by Σ_μ, and the second is the covariance among usage of the platforms on a given day, captured by Σ_e. Table 5 summarizes the estimates of the two covariance matrices. We observe a strong correlation between ESPN.com and ESPN3.com at the long-term level (cross-day posterior mean correlation = .795). Thus, heavy users of ESPN.com also tend to be heavy users of ESPN3.com, which is not surprising given that users access both platforms with the same type of device; thus, all users who access ESPN.com have the ability to access ESPN3.com. (Note that this correlation estimate is only for active users, because the spike at zero "absorbs" nonusers; the estimate of the correlation would have been higher if we had included the inactive users.) Indeed, knowing that there are some users with a high propensity to use ESPN.com but not ESPN3.com suggests a relatively easy opportunity for ESPN to expand viewership. Notably, we do not find a correlation between ESPN.com and ESPN3.com at the daily level (correlation = –.030). Using ESPN.com on a given day does not seem to be related to an increase in streaming video usage on ESPN3.com that same day.

The relationship with ESPN's mobile platform is significantly different and is of great business importance given the recent investments that ESPN (and many other media companies) has made in its mobile platforms. The mobile channel does not show any significant long-term or daily correlations with any other platforms. The only correlation that is directionally negative is the correlation between ESPN.com and ESPN Mobile at –.153 (2.5 percentile = –.244, 97.5 percentile = .512). This suggests that mobile usage is not cannibalizing usage of the other platforms (except, perhaps, ESPN.com) and is incremental.

Finally, we are able to estimate the within-day correlations between television and the other three platforms. As Table 5 indicates, the posterior intervals for all the correlations between television and the other platforms contain zero, suggesting that television usage on a given day is neither positively nor negatively correlated with the use of the digital channels. Notably, the posterior mean correlation between ESPN3.com and television is –.140 (2.5th percentile =

–.508, 97.5th percentile = .269), suggesting (directionally) that ESPN3.com and television do compete weakly with each other. This is consistent with the observation that television and ESPN3.com offer similar content (i.e., video of full games). By monitoring this parameter over time, as more data is accumulated, ESPN can keep better track of the relationship between ESPN3.com and television, an issue of key business importance.

In summary, we find no significant negative correlations between these four channels, suggesting that ESPN's content distribution platforms are not at saturation and that new platforms represent an opportunity to generate incremental reach. This is consistent with ESPN's belief that new platforms do not compete with the old but instead enable users to consume media at times that they previously could not. In addition, our finding that the mobile platform seems to provide incremental reach but is still not the most popular platform is consistent with ESPN's philosophy that users will choose "the best screen available at a given time" (Enoch 2009). We next provide a brief interpretation of the effects of the tournament content.

Tournament effects. As we described previously, the tournament effects include dummy variables for (1) whether a given day was on a weekend, (2) the number of games played, (3) the number of teams that must "win or go home" on a given day, (4) whether the U.S. team played, (5) whether one of three culturally significant teams (England, Australia, or Mexico) played, and (6) whether one of the three top-ranked teams (Spain, Brazil, or the Netherlands) played. Table 6 gives the posterior summaries for these coefficients.

Our results are consistent with the common notion in practice that, on weekends, people are more likely to watch television (β_{14} = .852) and less likely to go online (β_{11} = –.475 and β_{12} = –.396). (These parameters correspond to users being 2.3 times as likely to watch television on the weekend and about .6 times as likely to go online on the weekend.) However, we find no weekend effect for the mobile platform (β_{13} = .035). This provides important insight for ESPN planners; it seems that unlike the other media platforms, ESPN Mobile is equally accessible and used on both weekends and weekdays. Although we can only speculate on how consumers will use mobile in the future, this lack of a day-of-week effect suggests that media plans for the mobile platform will be different from those for television and online.

Turning to the tournament content itself, we observe a sensible significant effect for the number of games played on a given day. When there are a large number of games, ESPN.com becomes a more attractive platform, whereas the other platforms are relatively unaffected. We do not find effects on any of the platforms for the number of teams that must "win or go home." That is, we do not observe any evidence that "clincher" games attract more viewership on any of the platforms.[9]

With regard to our set of dummy variables for which teams are playing, we find that all platforms, particularly the mobile and television platforms, are more popular when the U.S. team is playing. The estimated parameters indicate users are 2.0 times as likely to access mobile content when the U.S. team is playing and 1.7 times as likely to watch television. Notably (and perhaps surprising to non–U.S. soccer fans), we find very weak (but positive) effects when a top team (Spain, Brazil, or the Netherlands) is playing, suggesting that the American audience we observe is more interested in the U.S. team than these top-rated soccer teams. For the variable that measures the aforementioned culturally significant teams, we find a weak positive effect for ESPN.com but not the other platforms.

Finally, we should note that we are able to achieve good fit with a relatively simple set of covariates describing how many games are being played and who is playing. Note that there are no covariates that describe the arc of the tournament (e.g., no dummies for the group stage vs. the knockout stage, the final game). Although we remain far from a complete theory of what makes a game attractive to watch on a particular platform, we note that we are able to capture the aggregate viewership (see Figure 2) with a relatively parsimonious set of covariates.

Table 6

ESTIMATED MODEL PARAMETERS: TOURNAMENT COVARIATES (β_K)

Parameters	*Mean*	*2.5th Percentile*	*97.5th Percentile*
Weekend			
ESPN.com	**–.475**	–.712	–.250
ESPN3.com	**–.396**	–.628	–.163
ESPN Mobile	.035	–.294	.328
Television	**.852**	.639	1.082
Number of Games			
ESPN.com	**1.370**	.842	1.980
ESPN3.com	–.231	–.842	.472
ESPN Mobile	.317	–.388	1.012
Television	.286	–.190	.705
Number of Teams That Must "Win or Go Home"			
ESPN.com	.174	–.304	.707
ESPN3.com	.095	–.495	.667
ESPN Mobile	–.145	–.817	.566
Television	.081	–.270	.492
U.S. Team Playing			
ESPN.com	.338	–.098	.821
ESPN3.com	.328	–.081	.741
ESPN Mobile	**.699**	.213	1.186
Television	**.523**	.227	.935
Canada, Australia, or Mexico Playing			
ESPN.com	.350	–.054	.726
ESPN3.com	.028	–.502	.444
ESPN Mobile	.075	–.425	.583
Television	.037	–.309	.415
Top Team Playing			
ESPN.com	.211	–.118	.557
ESPN3.com	.134	–.314	.563
ESPN Mobile	.073	–.375	.549
Television	.207	–.116	.544

Notes: When the posterior interval does not contain zero, the posterior mean appears in boldface.

Forecasting Alternative Media Plans

In this subsection, we present forecasts for two alternative media plans[10] that ESPN could have used instead of the

[9]We thank an anonymous reviewer for the suggestion to include this variable and report the null finding.

[10]We note that these plans are illustrative only and do not represent media plans that ESPN has been or might be considering.

large-scale "ESPN XP" program, which provided coverage for every game on all four platforms. We constructed the forecasts in the same way we computed the posterior predictions used in model assessment. We generated 100 posterior predictive data sets using 100 sets of parameters randomly sampled from the posterior draws obtained from the MCMC sampler. These predictions were based on an alternative set of covariate values describing an alternative media plan (for more detail, see the Web Appendix at www.marketingpower.com/jmr_webappendix).

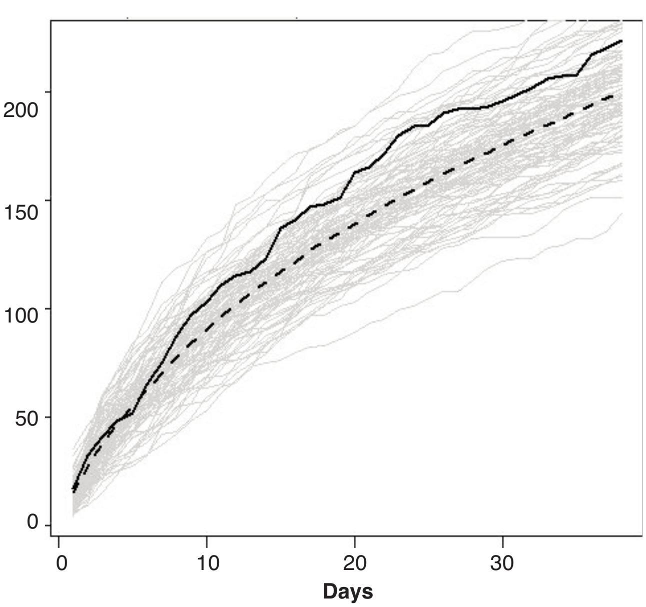
First, relating to the key business question of whether it is valuable to invest in coverage on the mobile platform, we created a scenario in which mobile coverage for the tournament was withdrawn entirely, leaving the coverage on the other platforms as it was. Figure 6 shows the predicted cumulative reach for the mobile channel (dotted line), which is somewhat lower than the actual reach when there was full mobile coverage (solid line). Consequently, mobile coverage of soccer does affect the mobile channel; however, when we observe reach across all the channels, we find no predicted drop in cumulative reach at the end of the tournament for all channels combined (49.7% forecast vs. 49.7% actual). This suggests that, at least in 2010, providing mobile coverage did not have a substantial impact on the total number of people who watched the tournament, most likely because overall mobile viewership was extremely low and thus did not contribute a great deal to the overall reach. Those who would have watched mobile were it available were also watching on other platforms. So, despite the finding that the mobile platform does not seem to be cannibalizing the other platforms (as evidenced by the estimates of the correlations between platforms) and is providing some incremental reach on some days, it does not seem to be providing a great deal of incremental cumulative reach over the course of the entire tournament.

Our second forecast represents a compromise scenario, in which ESPN provides mobile coverage only on days when the popular U.S. team is playing. Figure 7 plots the predicted cumulative reach for mobile had there only been coverage on those four days that the U.S. team played. We find that predictive cumulative reach is not substantially affected when mobile coverage is reduced; when comparing the prediction (dotted line) with the actual cumulative reach (solid line), we observe that the predicted reach is only slightly lower when the mobile coverage is reduced (and well within the band of prediction error). In total, we predict that 11.0% of the 2,000 users would have watched the mobile platform at all during the tournament compared with the 11.9% we observed in the actual data in which there was mobile coverage for all games in the tournament. (In contrast, when we forecast what would have happened had television coverage been reduced just to the days when the United States was playing, we find that reach for television is predicted to be substantially reduced.) This suggests that this compromise plan could have allowed ESPN to achieve the same reach for the mobile platform at a potentially lower cost by providing ESPN Mobile platform coverage only on days with the most interesting tournament content.

The forecasts presented here represent a small fraction of the types of forecasts that such a model can perform. Other forecasts that might be of interest include forecasting viewership for different outcomes of the tournament (e.g., "What

Figure 6

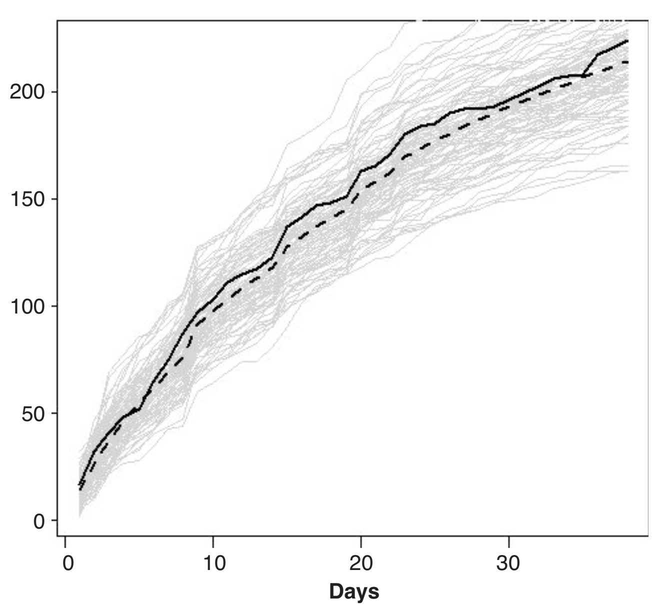
CUMULATIVE REACH FOR THE MOBILE PLATFORM HAD THERE BEEN NO MOBILE COVERAGE



Notes: We draw the forecast for cumulative reach and compare it with the actual statistic (computed directly from the sample), drawn with a solid line. To show the forecast uncertainty, we also draw the prediction for each of 100 random draws from the posterior with gray lines.

Figure 7

CUMULATIVE REACH FOR MOBILE HAD THERE BEEN MOBILE COVERAGE ONLY FOR U.S. GAMES



Notes: We draw the forecast for cumulative reach and compare it with the actual statistic (computed directly from the sample), drawn with a solid line. To show the forecast uncertainty, we also draw the prediction for each of 100 random draws from the posterior with gray lines.

if the U.S. team made it to the finals?"); however, this would require integration over the distribution of the set of potential tournament outcomes. If ESPN also had some influence over the tournament structure (which may be more likely for U.S.-based tournaments than the World Cup), researchers could use the model to predict media consumption impact due to changes in the tournament schedule (e.g., "What if all U.S. games were held on weekends instead of weekdays?"). All these forecasts enable us to assess the impact of media-planning decisions on the key economically meaningful outcomes: multiplatform reach and exposure.

DISCUSSION

As the number of media platforms proliferates, complicating planning problems for media companies, it is also important to appreciate the opportunities that the newer digital platforms offer. The rich, granular data sets that emerge from digital platforms provide media companies an unprecedented opportunity to track and model the behavior of individual users over time, without extraordinary data collection efforts. Researchers can use the resulting data to help better understand consumers and to improve business practice by modeling the interplay between platforms, as described here.

Unfortunately, the same is not true for traditional media platforms, and companies that aim to understand the relationship between traditional and digital media consumption have, until now, been forced to resort to expensive outside data collection (i.e., multiplatform media panels). We propose an alternative approach that combines readily available aggregate data from the traditional platforms with the individual-level data that digital platforms produce automatically. Companies can use the model we developed to assess cannibalization among platforms and to forecast media reach for alternative multiplatform media plans. Our data-fusion approach presents a way forward for analysts who have long been thwarted by differing levels of aggregation across platforms.

We expect researchers to apply this modeling framework more generally to other data structures that describe people's use of multiple platforms or channels over time. Data sets with a mixed structure, such as the one we present here, abound. For example, many multiplatform retailers collect data on purchase incidence for individual customers within their direct retail platforms but only have access to aggregate purchase data (sales-by-day-by-product) for brick-and-mortar stores. This is the same structure as the media data we discuss, and other researchers could analyze them within the same framework.

There are also several similar and related data structures that scholars could address using the Bayesian approach we describe here. For example, others might consider developing a similar approach that enables the fusion of digital media data for different groups of users for situations in which users cannot be tracked across platforms (However, this may require stronger assumptions about the relationships between platforms than we require here.) It may also be possible to extend our approach beyond media consumption tracking to advertising effectiveness measurement. Many companies record direct marketing events for individual customers and purchase events for those same customers but only have aggregate data on the number of mass media advertising exposures for those same customers. This structure is difficult to handle in individual-level advertising response models (e.g., Braun and Moe 2012). However, as Musalem, Bradlow, and Raju (2008) show, aggregation in the independent variables of a model can be handled using the same data augmentation strategy we used here to address aggregation in the dependent variables. Thus, companies could use data augmentation to estimate simultaneously which users were exposed to mass advertising and the effect of advertising on purchase. This application would be highly relevant to several online retailers (e.g., eBay, Expedia, eTrade, Overstock.com) that observe purchases for individual account holders and also purchase a significant quantity of television advertising.

Although the descriptive model we present here represents a practical tool for media companies to gauge and monitor multiplatform media consumption, we acknowledge that it does not provide a theoretical explanation for why media users choose to consume content on each platform. As data on media consumption become more widely available, we encourage researchers to propose theories on how users choose what to watch—that is, economic models for how people choose a set of platforms from which to consume content or behavioral models that describe how people attend to media content over time. As such data become more widely available, researchers will have more opportunities to test these types of theories empirically.

We conclude with some lessons learned from this practical case study and, in particular, from our experience working with potentially rich but messy digital media data sets. First, marketers should become much more connected with the information systems/computer science community. The ability to handle large databases and construct easily accessible data sets is not in many marketing researchers' skill sets, but it will need to be going forward. Second, there are many "wish list" items companies hope for, such as linking to advertisement data, linking to click-throughs on ads, real-time data, and more. Although these are nice ideas conceptually, there are practical limitations to what data we can actually obtain. Yet rather than "retreating," our experience suggests getting the best and most reliable data possible. In the rich tradition of applied research, we urge scholars to model the data "as it lies," and not as they might dream it to be.

REFERENCES

Albuquerque, Paulo and Bart J. Bronnenberg (2009), "Estimating Demand Heterogeneity Using Aggregated Data: An Application to the Frozen Pizza Category," *Marketing Science*, 28 (2), 356–72.

Ansari, Asim, Carl F. Mela, and Scott A. Neslin (2008), "Customer Channel Migration," *Journal of Marketing Research*, 45 (February), 60–76.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4), 841–90.

Braun, Micael and Wendy Moe (2012), "Online Advertising Modeling the Effects of Multiple Creatives and Individual Impression Histories," working paper, (accessed April 30, 2013), [available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1896486].

Chen, Yuxin and Sha Yang (2007), "Estimating Disaggregate Models Using Aggregate Data Through Augmentation of Individual Choice," *Journal of Marketing Research*, 44 (November), 613–21.

Chib, Siddhartha and Edward Greenberg (1998), "Analysis of Multivariate Probit Models," *Biometrika*, 85 (2), 347–61.

Crupi, Anthony (2010), "ESPN Fields Major Sports Research Initiative," *AdWeek*, (March 23), (accessed March 13, 2013), [available at http://www.adweek.com/news/television/espn-fields-major-sports-research-initiative-101910].

Danaher, Peter J. (2007), "Modeling Page Views Across Multiple Websites with an Application to Internet Reach and Frequency Prediction," *Marketing Science*, 26 (3), 422–37.

Enoch, Glenn (2009), "Seven Principles of Cross-Media Research," paper presented at the 2009 Joint Statistical Meetings, Washington, DC (August 1–6).

Franz, Gerhard (2000), "The Future of Multimedia Research," *International Journal of Market Research*, 42 (4), 459–72.

Gelman, Andrew, Xiao-Li Meng, and Hal S. Stern (1996), "Posterior Predictive Assessment of Model Fitness via Realized Discrepancies," *Statistica Sinica*, 6, 773–807.

——— and Donald B. Rubin (1992), "Inference from Iterative Simulation Using Multiple Sequences," *Statistical Science*, 7 (4), 457–511.

Gilula, Zvi, Robert E. McCulloch, and Peter E. Rossi (2006), "A Direct Approach to Data Fusion," *Journal of Marketing Research*, 43 (February), 73–83.

Glonek, G.F.V. and P. McCullagh (1995), "Multivariate Logistic Models," *Journal of the Royal Statistical Society: Series B (Methodological)*, 57 (3), 533–46.

Joo, Mingyu, Kenneth C. Wilbur, and Yi Zhu (2011), "Does Television Advertising Influence Online Search?" working paper, Whitman School of Management, Syracuse University, (accessed April 30, 2013), [available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1720713].

Little, Roderick J.A. and Donald B. Rubin (2002), *Statistical Analysis with Missing Data*. Hoboken, NJ: John Wiley & Sons.

Lin, Chen, Sriram Venkataraman, and Sandy D. Jap (2011), "Media Multiplexing Behavior: Implications for Targeting and Media Planning," working paper, (accessed March 13, 2013), [available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2153970].

Manchanda, Puneet, Asim Ansari, and Sunil Gupta (1999), "The 'Shopping Basket': A Model for Multicategory Purchase Incidence Decisions," *Marketing Science*, 18 (2), 95–114.

Morrison, Donald G. and David C. Schmittlein (1988), "Generalizing the NBD Model for Customer Purchases: What Are the Implications and Is It Worth the Effort?" *Journal of Business and Economic Statistics*, 6 (2), 145–59.

Musalem, Andres, Eric T. Bradlow, and Jagmohan S. Raju (2008), "Who's Got the Coupon? Estimating Consumer Preferences and Coupon Usage from Aggregate Information," *Journal of Marketing Research*, 45 (December), 715–30.

Naik, Prasad A. and Kay Peters (2009), "A Hierarchical Marketing Communications Model of Online and Offline Media Synergies," *Journal of Interactive Marketing*, 23 (4), 288–99.

——— and Kaylan Raman (2003), "Understanding the Impact of Synergy in Multimedia Communications," *Journal of Marketing Research*, 40 (November), 375–88.

Nevo, Aviv (2000), "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, 31 (3), 395–421.

Parker, Geoffrey G. and Marshall W. Van Alstyne (2005), "Two-Sided Network Effects: A Theory of Information Product Design," *Management Science*, 51 (10), 1494–504.

Rossi, Peter, Greg M. Allenby, and Robert E. McCulloch (2005), *Bayesian Statistics and Marketing*. New York: John Wiley & Sons.

Spiegelhalter, David J., Andrew Thomas, and Nicola G. Best (1999), *WinBUGS Version 1.2 User Manual*. Cambridge, UK: MRC Biostatistics Unit.

Tanner, Martin A. and Wing Hung Wong (1987), "The Calculation of Posterior Distributions by Data Augmentation," *Journal of the American Statistical Association*, 82 (398), 528–54.

Zheng, Zhiqiang, Peter Fader, and Balaji Padmanabhan (2012), "From Business Intelligence to Competitive Intelligence: Inferring Competitive Measures Using Augmented Site-Centric Data," *Information Systems Research*, 23 (3), 698–720.