International Journal of Research in Marketing 34 (2017) 761–779

Contents lists available at ScienceDirect

IJRM
International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Metrics unreliability and marketing overspending☆

Shrihari Sridhar [a,*], Prasad A. Naik [b], Ajay Kelkar [c]

[a] Center for Executive Development Professor, Mays Business School, Texas A&M University, United States
[b] Graduate School of Management, University of California Davis, United States
[c] Hansa Cequity Pvt. Ltd., Mumbai 400070, India

ARTICLE INFO

Article history:
First received on September 15, 2014 and was under review for 17 months
Available online 6 September 2017

Senior Editor: Koen H. Pauwels

Keywords:
Dynamic models
Measurement noise
Optimal budget and allocations
Kalman filter estimation
Emerging markets

ABSTRACT

The adverse consequences of measurement unreliability on statistical issues (e.g., inconsistency, attenuation bias) are well known. Yet there exists sparse literature, if any, on how unreliable metrics affect strategic marketing decisions: optimal marketing budget, its optimal allocation to advertising and promotions, and overspending. Consequently, researchers and managers do not know: How to estimate dynamic demand models using unreliable data? How to optimally combine multiple noisy and biased metrics? How to optimally set the total marketing budget and optimally allocate it to advertising and promotions activities using unreliable sales metrics?
To answer these open questions, first, based on Kalman filtering theory, we show how to estimate and infer dynamic demand models using unreliable sales metrics. Then, we furnish evidence of significant measurement noise in both retail audit and company's internal data to track brand sales. We replicate these results across six largest political regions in the emerging Indian markets for a major hair care brand. Next, we analytically derive the optimal weights to combine noisy and biased metrics to infer the latent demand. This result uncovers a counter-intuitive insight that two independent noisy metrics are better than one even when the second metric is noisier. In other words, a composite metric serves as noise reduction device as it is more reliable than individual noisy metrics. Subsequently, we derive closed-form expressions for the optimal budget and its optimal allocation to advertising and promotions activities in the presence of unreliable sales metrics. Based on these results, we discover that marketing overspending increase as metrics unreliability increases. Furthermore, overconfidence —the presumption that the metrics are reliable— leads to overspending on advertising and promotions. Managers should reduce advertising and promotional spending when sales metrics are noisy. Finally, we provide a simple correction factor that managers can use to eliminate overspending.

☆ We appreciate the comments and encouragement from Bill Bean, Worldwide Director, Colgate-Palmolive Company; David Ciancio, Managing Director, Customer Strategy and Engagement at Dunnhumby Inc.; Kevin Richardson, Senior Vice President Consulting Services at Nielsen; Harpreet Singh, CEO, Kvantum Inc.; M. G. Parameswaran, CEO, Draft FCB Ulka, India.We also thank the seminar participants at the Darden Business School, Goethe-Universitaet Frankfurt, Hass Business School, HEC Paris, Kellogg Business School, University of Mannheim, and the Wharton Business School for their helpful suggestions. We gratefully acknowledge the support of the participating company and Hansa Cequity for insights into managerial issues.
* Corresponding author.
E-mail addresses: ssridhar@mays.tamu.edu (S. Sridhar), panaik@ucdavis.edu (P.A. Naik), ajay.kelkar@cequitysolutions.com (A. Kelkar).

http://dx.doi.org/10.1016/j.ijresmar.2017.09.001
0167-8116/

1. Introduction

Virtually all metrics unreliably reflect their unobservable constructs, perhaps with the exception of time measured by atomic clocks or calendar weeks. In sciences, for example, metrics such as moisture, texture and pH measure the underlying "soil conditions" (Grace & Bollen, 2008). In economics, metrics like consumer confidence or gross domestic product reflect the state of an economy; or consumption and expenditure surveys help gauge poverty (e.g., Beegle, Weerdt, Friedman, & Gibson, 2010). In marketing, metrics such as sales, market share, cash flow, operating earnings, stock returns indicate a firm's value (e.g., Srinivasan & Hanssens, 2009). In branding, companies strive to win customers' hearts and minds and so they monitor mind-set metrics (e.g., awareness, liking, consideration, intention) to infer how customers think, feel and act (Bruce, Peters, & Naik, 2012; Pauwels, Erguncu, & Yildirim, 2013; Srinivasan, Vanhuele, & Pauwels, 2010). In advertising, eye-tracking data permits inference on viewers' angle of gaze, attention, information acquisition, and memory for brands (e.g., Wedel & Pieters, 2008). In sales, surveys of physicians indicate the information content used during sales call for detailing prescription drugs (Kappe & Stremerch, 2016). In retailing, consumer offtakes and competitors' brand sales are measured by using diary panels, scanner panels, or store audits to estimate marketing effectiveness (e.g., Nielsen, IRI, GfK).

Such metrics of awareness (Millward Brown), consumer confidence (Conference Board), customer satisfaction (ACSI), sales calls on physicians (IMS Health, Cegedim), consumption and expenditures (Bureau of Labor Statistics) are based on surveys, which are likely to be error-prone. Consumer offtakes measured by auditing stores are also contaminated by measurement errors, which we describe in detail in Section 4.1. Even bar code scanning for accurately tracking items sold at the stock-keeping unit (SKU) level when aggregated over dozens of SKU items in each store and hundreds of such stores dispersed geographically to obtain national brand sales yields various "estimates," depending on how SKUs are combined (e.g., weighted by volume, prices, or market shares).

Measurement errors manifest themselves in business-to-business (B2B) contexts as well. Companies report that 25% of B2B databases are inaccurate and 60% of them judged overall data quality as unreliable.[1] Hanssens (1998) states that accurate demand data are typically hard to come by, as most industries lack consolidated scanning services and instant demand feedback that are typical of the packaged goods sectors in advanced economies.

It is well known that the presence of measurement errors leads to adverse *statistical* consequences. For example, measurement errors render parameter estimates inconsistent and introduce attenuation bias (e.g., see Naik & Tsai, 2000). However, *marketing* consequences due to the presence of measurement errors are not known. Specifically, how does measurement noise impact optimal marketing budgeting, its optimal allocation to advertising and promotions, and marketing overspending? To this end, this paper develops a method to reduce measurement noise and discovers how metrics unreliability drives marketing overspending.

We begin by formulating a measurement model that incorporates bias and noise in each of the multiple metrics. Then, we derive the optimal weights to combine multiple metrics to infer the latent demand consistently. Next, we design a Kalman filter that controls for measurement errors, estimates the time-varying effectiveness of advertising and promotions, and quantifies the synergies between advertising and distribution as well as between promotions and distribution. We illustrate the application of the proposed method to a major hair care brand in India. Empirical results furnish evidence that measurement noise and bias in both the metrics and in each of the six regions are statistically significant at the 95% confidence level. Subsequently, we derive closed-form expressions for optimal advertising spending and optimal timing of promotions. We deduce new propositions that elucidate how metrics unreliability impacts the optimal budget and allocation as well as marketing overspending. Specifically, marketing overspending increases as metrics unreliability increases. This finding not only underscores the importance of metrics unreliability, but also incentivizes managers to reduce measurement noise. We also learn that overconfidence —the presumption that the metrics are reliable when they are not— drives overspending on advertising and promotions. This impact is asymmetric, with more overspending on advertising than on promotions. Finally, we derive a "correction factor" which offers a constructive approach for managers to eliminate overspending.

In sum, this paper is the first one to make the following original contributions. First, methodologically, we derive the optimal combination of unreliable metrics and incorporate it in the proposed estimation method. Second, empirically, we establish that metrics are indeed noisy and quantify the magnitude and heterogeneity across six markets in India. Third, theoretically, we discover how low versus high noise levels affect marketing overspending and its asymmetric effect on advertising and promotional overspending.

The rest of the paper is organized as follows. We first review the extant literature to establish gaps. We then formulate the measurement model to derive the optimal metrics combination to reduce noise. We then describe the data, estimation, and results. Subsequently, we derive new propositions on the effects of measurement noise on the optimal marketing budget, optimal allocation, and marketing overspending. We close by discussing the implications for managers and researchers.

[1] http://www.funnelholic.com/2013/09/16/the-problem-of-dirty-data-and-why-every-sales-and-marketing-leader-should-care/.

2. Literature review

2.1. Previous marketing-mix models using noisy sales data

We reviewed the marketing literature and found 166 studies using noisy sales metrics in marketing mix models. Of which, 89 studies included advertising effects (54%), 67 price effects (40%), 52 promotion effects (31%), and 19 distribution effects (11%). The sum exceeds 100% because multiple activities appear in some studies. We found various metrics for the dependent variable such as cross-sectional usage surveys of consumers, longitudinal consumer diary panel, shipments data from manufacturer, warehouse withdrawal (SAMI), and retail audit of order quantity and consumer offtakes. Assael (1967, p. 401) shows that consumer panel, retail audits, and telephonic survey are each individually unreliable (also see Wind & Lerner, 1979). Kuehn, McGuire, and Weiss (1966) conjectured that if noisy retail audit data are used to measure consumer demand, parameter estimates are possibly biased. But they did not estimate the magnitude of bias, which Shoemaker and Pringle (1980, p. 91) quantify and conclude that,

> "The results of several simulations suggest that this practice can lead to parameter estimates that are biased by as much as 22%. A bias of this magnitude could seriously affect conclusions reached by researchers and lead managers to make incorrect marketing decisions."

Despite this knowledge and concern, *none of the studies controlled for measurement noise in the sales metrics*. Hence, we formally adopt the errors-in-variables framework when estimating the marketing-mix effects.

Most studies did not use multiple sales metrics in the same study. Only 10 studies, which we list in Table 1, used multiple metrics. Leeflang and Olivier (1985) show that store audit data differ substantially from internal secondary data every month. Not unlike Shoemaker and Pringle (1980), they also show that.

> "… these deviations may lead to a large bias in parameter estimation, implying wrong indications for the use of marketing decision variables."

Table 1
Literature review.

Study	Results	Sales data sources	Lagged dependent variable?	Noise extracted in sales metrics?	Optimally combined sales metrics?
Assael (1967)	Show how three methods of determining market share (Consumer panel, retail audits and telephonic recall) are all individually unreliable.	Consumer panel, Retail store audit, consumer telephonic recall.	NA	No	No
Roshwalb (1970)	Show two different ways of performing retail audits to mitigate errors. Analytical paper with no data.	Retail audit measures from two methods (simulated).	NA	No	No
Brown (1973)	Show that changes in factory sales can be used to measure impact of marketing when inventory changes are known.	SAMI (selected area marketing information) data, factory sales data.	No	No	No
Wind and Lerner (1979)	Show discrepancies in consumer's purchase report (survey) and dairy panel (recording) entries and suggest qualitative ways to reduce discrepancies.	Surveys of consumer purchases, consumer dairy panel data.	NA	NA	NA
Shoemaker and Pringle (1980)	Use weekly advertising data to simulate daily sales, and correct bias in bi-monthly retail audit data.	Bi-monthly retail audit data, simulated daily sales data (from advertising data and model of ad/sales ratio).	Yes	No	No
Stanton and Tucci (1982)	Show that 24-hour recall and 2-day consumer diary panel data do not differ significantly.	Consumer recall, consumer diary panel.	No	No	No
Neslin and Shoemaker (1983)	Estimate effectiveness of consumer coupon promotions.	Consumer sales, retail order quantity.	No	No	No
Leeflang and Olivier (1985)	Secondary sales, retail audit and consumer panel data are all individually unreliable.	Secondary sales, retail audit, consumer panel	No	No	No
Abraham and Lodish (1993)	Estimate effectiveness of trade and consumer promotions.	Manufacturer shipments, consumer sales.	No	No	No
Blattberg and Levin (1987)	Estimate effectiveness of trade promotions.	Manufacturer shipments, consumer sales.	No	No	No
Present Study	Estimate marketing-mix effectiveness in emerging markets by optimally combining and denoising multiple unreliable sales metrics.	Secondary sales, RETAIL offtakes.	Yes	Yes	Yes

Moreover, they find that parameter estimates obtained from each metric differ significantly. So they formally test the hypothesis of equality of the corresponding parameters and conclusively reject it for all the six brands they analyzed except one. Thus they conclude that the various metrics "… give different indications of the use of marketing decision variables."

More importantly, *none of the studies in* Table 1 *combined multiple metrics*. When multiple noisy metrics are available, their weighted combination facilitates the reduction in variance. To understand this phenomenon, consider a simple example of two independent normal variables $X_j \sim N(\mu_j, \sigma^2)$, $j = 1$ and 2. Although the variance of each variable equals σ^2, the variance of the combined variable $0.5X_1 + 0.5X_2$ is $\sigma^2/2$, which is half as much as any one of them. This principle of variance reduction forms the basis for diversification of risky assets in portfolio management (Markowitz, 1952) or combination of forecasts in econometrics (Bates & Granger, 1969) or creation of antithetic variables in Monte Carlo studies (Morgan, 2005). In other words, we gain efficiency by combining multiple unreliable metrics. Hence, we formally derive the optimal weights to combine multiple metrics. Next, we briefly review the statistics literature on errors-in-variables.

2.2. Errors-in-variables literature

Given the adverse consequences of ignoring measurement errors (e.g., inconsistency, attenuation bias), a vast literature exists on the topic of errors-in-variables, and hence we review a few key studies. Adcock (1877, 1878) first posed the problem of errors in independent variables and proposed the method of orthogonal regression that minimizes orthogonally projected distance from data points to the regression line (instead of the vertical distances as in the standard regression). Pearson (1901) extended this idea to multiple independent variables. Lindley (1947) pioneered the use of maximum likelihood estimation for the errors-in-variables model. To tackle errors in both the independent and dependent variables, Jöreskog (1970) proposed the confirmatory factor analysis, where measurement equations incorporate the errors in both kinds of variables and the structural equations specify the relations among the latent factors. This approach is commonly used in marketing (e.g., see Grewal, Cote, & Baumgartner, 2004, Ramani & Kumar, 2008, Steenkamp & de Jong, 2010).

Despite the century of research on this topic, the above models do not apply to time series data with inter-temporal dependence in variables. Notable exceptions include Cai, Naik, and Tsai (2000), who proposed the denoised least squares estimator, where each noisy metric is individually de-noised by applying an appropriate wavelet transform before fitting the resulting denoised variables via time series models (see Naik & Tsai, 2000 for a marketing application). To jointly filter noise from multiple noisy metrics observed over time, Bruce et al. (2012) specify dynamic factor models, where multiple metrics reflect the latent factors that evolve over time. Dynamic factor models are special cases of the general state space models and the Kalman filter (see Harvey, 1994).

2.3. Implications for model development

Based on the above review, we identify the gaps in the extant literature. First, the sales metrics such as retail offtakes and secondary sales contain measurement noise, which induces biases in the estimated parameters. Second, the extant marketing models based on noisy sales metrics observed over time have not used errors-in-variables formulation to filter noise when estimating parameters. Third, the extant marketing models have not optimally combined the sales metrics to gain efficiency in parameter estimation. Last but not least, the extant literature does not present the derivation of the optimal budget and its allocation to advertising and promotions in the presence of unreliable metrics.

To fill these gaps, we next formulate the errors-in-variable model to control for measurement noise, optimally combine the multiple metrics to infer the latent demand, and formally derive the optimal advertising spending and promotion timing in the presence of unreliable metrics.

3. Model development

We first formulate a new marketing model that addresses the two challenges: controlling metrics unreliability and combining information from multiple metrics. Then we describe parameter estimation and robust inference.

3.1. Controlling metrics unreliability

We use errors-in-variables framework to model unreliable metrics, which accounts for measurement noise and bias in each metric. Let Y_{1t} and Y_{2t} denote the observed retail offtakes and secondary sales at time t, and ε_{jt} $(j = 1, 2)$ be the measurement errors in each metric. Because both metrics reflect the common demand, S_t, we model the errors in the metrics as follows:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \pi_1 S_t \\ \pi_2 S_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \tag{1}$$

where π_j measures the bias (in retail offtakes or secondary sales). The error vector $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ follows the bivariate normal with zero means and the error variances σ_j^2 $(j = 1, 2)$ are arranged diagonally in a matrix R. Eq. (1) thus formally incorporates measurement errors in both the metrics, whose presence can be ascertained by testing for statistical significance of the variances.

Marketing activities such as advertising and promotions drive the latent demand. Previous research (e.g., Hanssens, Parsons, & Schultz, 2001; Leone, 1995) shows that the carryover effect λ from the lagged demand drives the inter-temporal influence of marketing actions (e.g., the carryover effects from past advertising or promotions). Eq. (2) captures the effects of current marketing activities and the carryover effects from past marketing activities as follows:

$$S_t = \lambda S_{t-1} + \beta_{1t}\sqrt{u_t} + \beta_{2t}v_t + \eta_t, \tag{2}$$

where S_t is the latent consumer demand, λ is the carryover effect, (β_{1t},β_{2t}) are time-varying effectiveness of advertising and promotions, (u_t,v_t) are advertising spending and a promotions "on-off" indicator, respectively, and η_t denotes the normal errors in demand specification. The square root function captures the diminishing returns to advertising (see Simon & Arndt, 1980), which means the incremental sales from additional advertising diminish as spending levels increase.

The effectiveness of advertising and promotions are usually assumed to be constant over time. We relax this assumption for two reasons. First, constant effectiveness models imply constant optimal spending over time (see Naik & Raman, 2003). But actual spending varies over time, contradicting this predicted pattern. Hence, we relax the assumption by making the effectiveness parameters vary over time (e.g., Ataman, Van Heerde, & Mela, 2010). Because the exact nature of time variation is not known, recent studies specify random walk evolution (e.g., Kolsarici & Vakratsas, 2010), which parsimoniously captures non-monotonic dynamics.

Second, marketing in one medium may enhance the effectiveness of marketing in another medium. Consumers often see a brand's campaign in one medium, but forget the exact nature of the marketing efforts. When they see it again in a different medium, it reinforces their memory and they recall the campaign in the first medium. Thus, the marketing response in one medium enhances due to the presence of the second medium. Such reinforcement induces synergies between media (e.g., Naik & Raman, 2003; Narayanan, Desiraju, & Chintagunta, 2004). To capture such synergies between advertising and distribution as well as promotions and distribution, we extend the random walk model by including a non-zero drift to get $\beta_{i,t+1} = \beta_{i,t} + \gamma_i z_t + \eta_{it}$, where z_t denotes the distribution intensity. In other words, if a brand is more widely available, then its advertising and promotions are likely to be more effective.

We extend Eq. (2) by incorporating both the extensions as follows:

$$\begin{bmatrix} S_t \\ \beta_{1,t+1} \\ \beta_{2,t+1} \end{bmatrix} = \begin{bmatrix} \lambda & \sqrt{u_t} & v_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_{t-1} \\ \beta_{1,t} \\ \beta_{2,t} \end{bmatrix} + \begin{bmatrix} 0 \\ \gamma_1 z_t \\ \gamma_2 z_t \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \\ \eta_{3t} \end{bmatrix}, \tag{3}$$

where the error vector $\eta_t = (\eta_{1t},\eta_{2t},\eta_{3t})'$ follows the trivariate normal $N(0,Q)$ with zero means and the covariance matrix Q. The variance captures the unexplained portion of the variation in the true demand S_t.

Eq. (3) represents the transition equation in the state space framework (see Harvey, 1994), where $\alpha_t = (S_t,\beta_{1,t+1},\beta_{2,t+1})'$ is the state vector, $d_t = (0,\gamma_1 z_t,\gamma_2 z_t)'$ is the drift vector, and the matrix in Eq. (3) is called the transition matrix T_t. We link the state vector to the unreliable metrics $Y_t = (Y_{1t},Y_{2t})'$ in Eq. (1) via the observation equation:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \pi_1 & 0 & 0 \\ \pi_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} S_t \\ \beta_{1,t+1} \\ \beta_{2,t+1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \tag{4}$$

where we denote the link matrix in the above equation by H. The above discussion completes the model specification, which can be expressed compactly in the state-space form: $Y_t = H\alpha_t + \varepsilon_t$ (Eq. (4)) and $\alpha_t = T_t\alpha_{t-1} + d_t + \eta_t$ (Eq. (3)).

3.2. Combining multiple metrics optimally

The two metrics do not match exactly over time; that is, $Y_{1t} \neq Y_{2t}$ for various t. One way to combine the multiple metrics is to update the states proportional to the forecasting errors as follows:

$$\hat{\alpha}_t = \hat{\alpha}_{t|t-1} + K_t\left(Y_t - \hat{Y}_t\right), \tag{5}$$

where $\hat{\alpha}_{t|t-1} = E[\alpha_t|\Im_{t-1}]$, $\widehat{Y}_t = E[Y_t|\Im_{t-1}]$, $\Im_{t-1} = \{Y_1, \cdots, Y_{t-1}\}$ denotes the information set at time $(t-1)$, and K_t is a 3×2 time-varying matrix to be determined. The elements $\{k_{11,t},k_{12,t}\}$ in the matrix K_t are the weights placed on the forecasting errors to obtain $\hat{S}_t$.Similarly, using the other elements in second and third rows of the matrix K_t, Eq. (5) updates the estimates of advertising and promotions effectiveness.

We seek the optimal weighting matrix K_t^* such that the estimates $\hat{\alpha}_t$ are as close as possible to their true values α_t on average. Formally stated, let $\omega_t = \alpha_t - \hat{\alpha}_t$, so that the mean squared error is given by.

$$J_t = E\left[\left(S_t - \hat{S}_t\right)^2 + \left(\beta_{1,t+1} - \hat{\beta}_{1,t+1}\right)^2 + \left(\beta_{2,t+1} - \hat{\beta}_{2,t+1}\right)^2\right] = E\left[\omega_t'\omega_t\right] = E\left[Tr\left(\omega_t'\omega_t\right)\right] = Tr(E[\omega_t\omega'_t]) = Tr(P_t), \tag{6}$$

where the third equality follows by noting $\omega_t'\omega_t$ is a scalar; and the fourth one interchanges the trace and expectation operators and sums the diagonal of the matrix P_t. Next, we prove in Appendix A that.

$$P_t = (I - K_tH)P_{t|t-1}(I - K_tH)' + K_tRK_t', \tag{7}$$

where $P_{t|t-1} = Var[\alpha_t|\Im_{t-1}]$. Finally, to bring the estimates closest to the true values, we choose the matrix K_t that minimizes Eq. (6). Recalling that $\partial Tr(ABA') = 2AB\partial A$ for symmetric B, we obtain the first-order condition:

$$\frac{\partial J_t}{\partial K_t} = 2(I - K_t H)P_{t|t-1}(-H)' + 2K_t R. \tag{8}$$

By setting $\partial J_t/\partial K_t = 0$, we determine the optimal weighting matrix as follows:

$$\begin{aligned} &(I - K_t^* H)P_{t|t-1}H' = K_t^* R \\ &P_{t|t-1}H' = K_t^*\left(HP_{t|t-1}H' + R\right) \\ &K_t^* = P_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1} \end{aligned} \tag{9}$$

The last equality in Eq. (9) provides the optimal weights to combine the multiple metrics Y_t in Eq. (5). Appendix B derives the closed-form expressions given in

Proposition 1. *The optimal combination of the unreliable metrics Y_1 and Y_2 is given by $\hat{S} = k_1^* Y_1 + k_2^* Y_2$, where $k_1^* = \frac{\pi_1\phi_1}{1+\pi_1^2\phi_1+\pi_2^2\phi_2}$ and $k_2^* = \frac{\pi_2\phi_2}{1+\pi_1^2\phi_1+\pi_2^2\phi_2}$ are the optimal weights, $\phi_1 = \frac{\sigma_S^2}{\sigma_1^2}$ and $\phi_2 = \frac{\sigma_S^2}{\sigma_2^2}$ are the signal-to-noise ratios of the two metrics, and σ_S^2 denotes the mean squared error of latent demand.*

Proof. See Appendix B.

Before we discuss insights from this proposition, we note that the optimal weights vary over time because the gain matrix K_t^* in (9) is time varying, although we suppressed the time subscript here, but incorporate it in the next section on parameter estimation. Also, $\hat{S}$ is not an unbiased estimator of S as it exhibits "shrinkage" akin to Bayesian estimators, which balance the bias-variance tradeoff to increase efficiency by tolerating bias. An unbiased estimator necessarily sacrifices precision. Finally, we define the terms we use interchangeably, although they have distinct meanings. Specifically, for each metric j, measurement error refers to ϵ_j; measurement noise means σ_j^2; signal-to-noise ratio is $\phi_j = \sigma_S^2/\sigma_j^2$; and metric unreliability is noise-to-signal ratio, $\phi_j^{-1} = \frac{\sigma_j^2}{\sigma_S^2}$.

Four original insights emerge from the above proposition. First, the optimal weights depend on the two aspects of unreliability: signal-to-noise ratios $\phi_j$and biases π_j. As the signal-to-noise ratio for the metric improves, its weight increases in informing the true consumer demand. Furthermore, when π_j equals unity so that the relative bias disappears, the weights become symmetric functions of signal-to-noise ratios (i.e., $k_j^* = \frac{\phi_j}{1+\phi_1+\phi_2}$).

Second, the derivation of optimal weights does not require the assumption that the measurement errors are normally distributed. In other words, the weights given in Proposition 1 are optimal across *any* distribution of measurement errors with finite first two moments.

Third, *a composite metric serves as noise reduction device as it is more reliable than independent noisy metrics*. Researchers and managers can improve reliability by using multiple noisy metrics. Most researchers do not combine multiple noisy metrics (see Table 1). For example, Bass et al. (2007, p. 184) had data on two metrics of demand —minutes of call time and number of calls— but they did not combine them because the optimal weights were not known, as Proposition 1 was not available then.

Lastly, *even the inclusion of weak metrics improves reliability*. To understand this point, suppose a team of researchers has one reasonable metric and then they find another independent one that is twice as noisy. Should they combine the two metrics or just use the precise one? In other words, does the inclusion of a weak metric dilute the mix? Given that the weak metric is twice as noisy as the precise one, let's consider a convex combination in 2:1 proportion. Then, the variance of the composite metric is given by $Var(\frac{2}{3}Y_1 + \frac{1}{3}Y_2) = \frac{4}{9}\sigma^2 + \frac{1}{9}(2\sigma^2) = \frac{2}{3}\sigma^2$, which is smaller than the noise σ^2 in the more precise metric. In other words, an additional noisier metric injects new information, thus serving as a noise reduction device. Hence *the composite metric becomes more reliable by inclusion of an independent noisier metric*.

3.3. Parameter estimation and robust inference

To assess marketing effectiveness in the presence of unreliable sales metrics, we apply Eqs. (5) and (9) starting with the initial values α_0 and model parameters $(\lambda, \gamma_i, \pi, Q, R)$, whose values managers don't know when the models or markets are new. Hence we describe how to estimate parameters via the maximum-likelihood theory. Specifically, we first compute the log-likelihood function,

$$LL(\theta) = \Sigma_{t=1}^{T} Ln[p(Y_t|\Im_{t-1})], \tag{10}$$

where $p(\cdot|\cdot)$ denotes the conditional density of Y_t based on the metrics observed up to the previous period, $\Im_{t-1} = \{Y_1, \cdots, Y_{t-1}\}$. Then, using Eq. (4), we find the conditional mean $\hat{Y}_t = E[Y_t|\Im_{t-1}] = Ha_{t|t-1}$, so the innovation errors $(Y_t - \hat{Y}_t)$ are distributed with zero mean and the covariance matrix $F_t = HP_{t|t-1}H' + R$, where $(\hat{\alpha}_{t|t-1}, P_{t|t-1})$ are the conditional means and covariances of the "prior" state vector $\alpha_t \mid \Im_{t-1}$. We obtain its moments via Eqs. (3) and (4). Specifically, $\hat{\alpha}_{t|t-1} = T_t\hat{\alpha}_{t-1} + d_t$ and $P_{t|t-1} = T_t P_{t-1} T_t' + Q$, where $(\hat{\alpha}_{t-1}, P_{t-1})$ are the conditional means and covariances of the "posterior" state vector $\alpha_{t-1} \mid \Im_{t-1}$. After the new data arrives, that is, $\Im_t = Y_t \quad \Im_{t-1}$, we update the prior moments via Eqs. (5), (7), and (9). Then, ignoring the irrelevant constants, we recursively build the log-likelihood function,

$$LL(\theta) = -0.5\Sigma_{t=1}^{T}[Ln\left(\det(F_t)\right) + \left(Y_t - \hat{Y}_t\right)' F_t^{-1}\left(Y_t - \hat{Y}_t\right)\Big], \tag{11}$$

where det(·) denotes the determinant. For further details, see Appendix B in Naik, Mantrala, and Sawyer (1998).

Next, to estimate the parameter vector $\theta = (\lambda', \gamma', \pi', \text{diag}(Q), \text{diag}(R))'$ with p elements (as necessary for multiple regions), we maximize Eq. (11) to obtain,

$$\hat{\theta} = ArgMax\ LL(\theta). \tag{12}$$

Finally, to obtain the standard errors of $\hat{\theta}$, we take the square root of the diagonal of the inverse of the matrix:

$$\hat{C} = -\frac{\partial^2 LL(\theta)}{\partial\theta\partial\theta'}\bigg|_{\theta=\hat{\theta}}, \tag{13}$$

where the Hessian of $LL(\theta)$ is evaluated at the estimated values $\hat{\theta}$.

To make inferences robust to mis-specification errors (e.g., the functional relation specified between marketing spends and sales outcomes differ from the actual market reality), we compute the sandwich estimator (White, 1982):

$$\tilde{C} = \hat{C}^{-1}\hat{V}\hat{C}^{-1}, \tag{14}$$

where V is a $p \times p$ matrix of the gradients of the log-likelihood function; that is, $V = G'G$ and G is $T \times p$ matrix obtained by stacking the $1 \times p$ vector of the gradient of $LL(\theta)$ for each of the T observations. In correctly specified models, $C = V$ and so both the Eqs. (13) and (14) yield exactly the same standard errors (as they should); otherwise, we use the robust standard errors given by the square root of the diagonal elements of $C\sim$, which safeguards researchers from general mis-specification errors. The next section describes the data and the retail store auditing process that creates metrics unreliability, and then presents the estimation results.

4. Empirical analysis

4.1. Data

We analyze the marketing activities of a major brand of hair care in India. Due to nondisclosure agreements, we cannot divulge the brand's identity and the proprietary data. But the brand is well known and generates several million US dollars in annual revenues. It is distributed widely across urban, semi-urban, and rural regions. Over 130 million consumers are reached through its distribution network of over 3 million outlets in India. Given the vast coverage of the brand to retailers that lack bar code technology to track sales, the firm employs A. C. Nielsen Inc. to conduct field audits of a small sample of selected stores.

Nielsen's retail sales audit methodology entails two main activities: (a) shop census and sample determination every three years, and (b) shop audits and sales projection on a monthly basis. We next describe this process in detail. Activity (a) begins by creating 19 Nielsen regions from thirty five political states of India. These regions come from the four zones: North (Delhi, Punjab, Haryana, Rajasthan, Uttar Pradesh, Uttaranchal), West (Gujarat, Maharashtra/Goa, Madhya Pradesh, Chhattisgarh), South (Andhra Pradesh, Karnataka, Kerala, Tamil Nadu/Pondicherry), and East (Bihar, Jharkhand, West Bengal, Orissa, Assam & N.E). Across these 19 regions, the first stage sampling draws 462 urban towns and 1427 rural villages, which represent 8.9% and 0.24% of towns and villages, respectively. In each sampled town or village, stores dealing in products of interest are classified according to their type (Grocer, General Store, Chemist, Paan-Plus, Food Store, Cosmetic Store, Others) and an assigned score called dealer class value. Using weights in the parentheses, the dealer class value of a store is computed depending on its monthly turnover (52%), the number of fast moving consumer products it stocks (28%), the number of salespersons it employs (11%), and the number of cooling devices it contains (9%). Then, based on these two dimensions (store types by dealer class value), the second stage sampling selects 10,653 stores in urban towns and 5593 stores in rural villages, which totals to 16,246 stores nationwide. Even though most towns with population over a million and over one-half of all the districts in each region are included, this panel of shops represents a meager 0.21% of the retail universe of 7.87 million stores.

Activity (b) involves sending field agents to inspect the panel of shops between 12th and 28th of every month. They record the monthly inventory (validated by taking pictures) and new purchases (allegedly verified by billing information) across various product-packs from 112 product categories. Using this data, Nielsen computes the monthly retail offtakes based on the equation: retail offtakes equals opening inventory plus new purchases minus closing inventory. Based on the census data, they finally project the offtakes from the panel stores to the Nielsen regions and national levels, and report it within 45 days to the companies subscribing these services.

We spoke with several industry leaders about the unreliability of Nielsen's retail offtakes metric. Mr. M. G. Parameswaran, CEO of Draft FCB Ulka, summarizes the situation: “The challenge is not just in the small sample sizes but the inter-city, inter-region variations. In every product category there are regional variations. Often this is not captured accurately by the audits.” Besides the small sample of the retail panel and the non-representativeness of inter-region and inter-city variations across product categories, many other errors creep in because of the manual nature of the field work, multiple purchases after the auditors' visits, untracked inventory in offsite locations and home attics, unbilled new purchases to avoid taxes and to under-report revenues, lack of representation of modern trade, slow updates of shop census and stores' panel relative to changing market conditions, among others.

The resulting problem of measurement noise in retail offtakes is so severe that major companies such as Unilever, Godrej, and Dabur went public with their dissatisfaction, according to Balakrishnan (2011), who writes in *Economic Times* that Unilever considered dropping Nielsen's services and Mr. Paul Polman, the global CEO of Unilever, expressed reservations about the accuracy of the reported offtakes. To remedy the situation, Mr. Piyush Mathur, CEO of Nielsen India, admits that the retail audit is a "constant work-in-progress" and proposes to expand the panel from 16,246 to 24,000 stores. Nonetheless, Mr. Dalip Sehgal, Managing Director of Godrej Consumers Products, says that "Nielsen would have us believe that increasing the sample would resolve issues. But past records tell a different story. Samples were increased in 2006, and yet, reliability of data is no better, and in some cases, even worse" (Balakrishnan, 2011 in *Economic Times*).

Despite this unreliability, large consumer product companies subscribe to Nielsen's data. What justifies their decisions to continue the subscriptions to Nielsen's unreliable data? Proposition 1 does. Albeit unreliable, Neilsen's retail offtakes provides some corroboration and new information —akin to a "second opinion"— to aid their understanding based on internal secondary sales data, which is not a perfect metric either. As noted in the main implication of Proposition 1, the reliability of the composite metric improves when two noisy metrics from independent sources are used rather than just one, even if the second metric is substantially biased and noisy.

4.2. Variables

The two sales metrics are retail offtakes (i.e., quantity sold to consumers) and the company's secondary sales (i.e., quantity sold to retailers). The brand is available in three pack sizes (100 ml, 200 ml, 500 ml) in multiple flavors with a total of 18 stock-keeping units, and so we analyze the total volume in kiloliters. Advertising and promotions vary over time, while prices remained nearly constant over time and identical across regions (and hence cannot be included in the model). Advertising data include the total GRPs in national and cable television, and promotion data indicates the timing of promotions. We augment this information with distribution intensity over time, i.e. the percentage of retailers who carried the brand. Accordingly, we can test whether distribution creates synergies with advertising and promotions. In other words, in the presence of synergy, advertising (or promotion) effectiveness should be enhanced due to wider brand availability: the greater the penetration, the more effective the advertising (or promotions). We present results from the largest six political regions, covering over 80% of the total sales, spanning the breadth of the country: Andhra Pradesh, Gujarat, Karnataka, Maharashtra, Tamil Nadu, and Uttar Pradesh. All data are observed over 33 months starting from April 2008 to December 2010. Fig. 1 provides a plot of the two sales metrics in each region, and Table 2 presents the descriptive statistics.

For each region, we estimate the proposed model in Eqs. (3) and (4) by applying the estimation approach described in Section 3.3. In addition, we account for potential endogeneity in advertising and promotions by using an instrumental variables approach (e.g., Aravindakshan, Peters, & Naik, 2012, Bronnenberg & Mahajan, 2001). We predict each region's advertising spending using spending in all other regions, and use this predicted spending as the regressor for advertising. Similarly, we predict each region's promotional timing using the seasonality index from other political regions and other products in the category, and use this predicted promotional timing as the regressor for promotions. We next describe the empirical results.

4.2.1. Fit and forecasts

Table 3 shows the fit and forecast for all six regions. As Table 3 shows, the model fits the data from all six regions satisfactorily. For example, in Maharashtra, the fit for retail offtakes (MAPE = 10.33%) is better than that for secondary sales (MAPE = 16.02%). Similarly, the out-of-sample forecasts are satisfactory. Specifically, we estimate the model using 28 observations and evaluate the forecast errors based on the last 5 observations in the holdout sample. For example, in Tamil Nadu, the out-of-sample for secondary sales (MAPE = 11.81%) is better than that for retail offtakes (MAPE = 14.56%). We next describe the parameter estimates.

4.3. Unreliability estimates

We focus on the estimates of measurement noise and relative bias. Table 4 presents the parameter estimates and robust t-values for all the six regions. First, the measurement noise in retail offtakes is large and significant in all the six regions (Andhra Pradesh: $\sigma_1 = 11.97, t = 4.91$, Gujarat: $\sigma_1 = 8.14, t = 8.72$, Karnataka: $\sigma_1 = 11.72, t = 6.29$, Maharashtra: $\sigma_1 = 18.41, t = 13.48$, Tamil Nadu: $\sigma_1 = 6.23, t = 7.86$, Uttar Pradesh: $\sigma_1 = 7.05, t = 4.70$).

Second, the measurement noise in secondary sales also is large and significant in all six regions (Andhra Pradesh: $\sigma_2 = 27.60, t = 6.99$, Gujarat: $\sigma_2 = 22.46, t = 6.12$, Karnataka: $\sigma_2 = 17.05, t = 5.74$, Maharashtra: $\sigma_2 = 67.64, t = 5.85$, Tamil Nadu: $\sigma_2 = 5.14, t = 5.57$, Uttar Pradesh: $\sigma_2 = 5.21, t = 2.56$). Thus, systematically across all regions, these results furnish strong empirical evidence for metrics unreliability.

Third, the bias in both the metrics is large and significant in all the six regions (see estimates and t-values of $(\hat{\pi}_1, \hat{\pi}_2)$ in Table 4). Ignoring this presence of unreliability renders all parameter estimates inconsistent (Naik & Tsai, 2000). In other words,

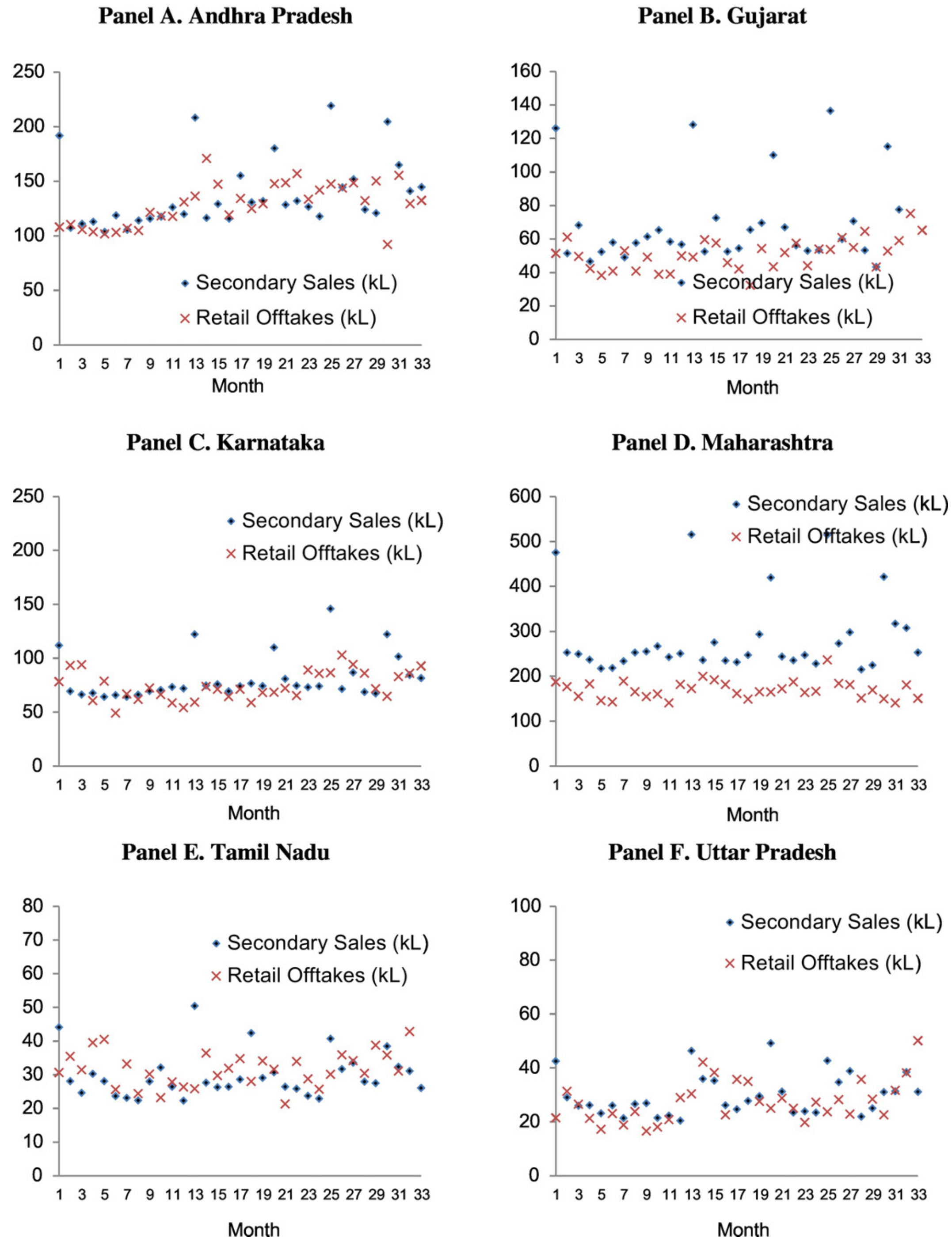

Fig. 1. Retail offtakes and secondary sales across regions.

managers will estimate parameters of marketing models inaccurately even if the sample size were asymptotically large. In contrast, the proposed approach circumvents this problem by filtering out the measurement noise via Eq. (1).

4.4. Carryover and synergy effects

The consumer demand exhibits strong carryover effects in all the six regions —it is large and significant with the median value of 0.91 and ranges from 0.87 to 0.94.

Table 2
Descriptive statistics.

	Andhra Pradesh	Gujarat	Karnataka	Maharashtra	Tamil Nadu	Uttar Pradesh
Retail offtakes (average), (kilo-liters)	128.86	50.70	74.24	169.74	31.93	27.48
Retail offtakes (standard deviation)	19.52	9.34	13.46	20.20	5.58	7.70
Secondary sales (average), (kilo-liters)	137.25	68.43	80.94	284.29	29.77	29.81
Secondary sales (standard deviation)	31.14	24.99	19.91	84.84	6.67	7.64
Advertising GRPs (average)	290.73	1571.98	252.35	325.27	1740.38	1281.39
Advertising GRPs (standard deviation)	290.42	1476.83	249.72	224.80	1642.92	1169.81
% Promotion on-off (average)	21.2	21.2	21.2	21.2	21.2	21.2
% Promotion on-off (standard deviation)	41.5	41.5	41.5	41.5	41.5	41.5
% Retailers carrying brand (average)	61.2	27.82	49.6	64.7	33.32	47.62
% Retailers carrying brand (standard deviation)	1.23	1.35	2.00	2.16	2.28	3.36

How does distribution impact the effectiveness of advertising and promotions? Hanssens (1998) says that, "Distribution would ideally be measured by number or percentage of retail outlets carrying the product, … but such data are not available." Fortunately, we do measure the number of retail stores carrying the firm's brand. Then, we find that advertising effectiveness significantly increases with brand availability in Andhra Pradesh ($\gamma_1 = 0.06$),Karnataka ($\gamma_1 = 0.03$), and Maharashtra ($\gamma_1 = 0.001$). We also furnish empirical evidence for synergy between distribution and promotions. Specifically, promotional effectiveness increases as the distribution intensity increases in Gujarat ($\gamma_2 = 3.61$), Karnataka ($\gamma_2 = 3.95$), Maharashtra ($\gamma_2 = 0.35$), Tamil Nadu ($\gamma_2 = 4.81$),and Uttar Pradesh ($\gamma_2 = 3.35$). These results, new to the literature, empirically ground our understanding of the distribution-driven synergistic effects.

4.5. Marketing elasticities

To compare the effects of advertising and promotions, we compute their elasticity, which is a dimensionless quantity and hence comparable across variables with different measurement units. Recall that elasticity means 1% change in advertising (or promotions) results in ω (or υ) percentage change in consumer demand. Denoting $(\overline{z_i}, \overline{u_i}, \overline{S_i})$ as the mean values of distribution, advertising and consumer demand, respectively, we derive from Eq. (3) the advertising elasticity $\omega_i = (0.5\gamma_{i1}\overline{z_j}\sqrt{\overline{u_i}})/\overline{S_i}$ and the promotions elasticity $\upsilon_i = (\gamma_{i2}\overline{z_i})/\overline{S_i}$ for the region *i*.

Table 5 presents advertising and promotions elasticities. Across the six regions, the median advertising elasticity is 0.0503, which means the latent market demand increases by about 5% if advertising doubles. This estimate is smaller than the average ad elasticity of 0.10 (i.e., the demand increases by 10% if advertising doubles) in meta-analysis by Sethuraman, Tellis, and Briesch (2011). Three factors explain our lower ad elasticity. Our setting represents a non-durable good (hair care oil), in a mature product lifecycle stage, with moderate advertising support. Sethuraman et al. (2011) show that ad elasticity tends to be higher (1) for durable goods than nondurable goods, (2) in the early stage than the mature stage, and (3) when products are heavily advertised. Thus, the smaller ad elasticity corresponds with the opposite market characteristics.

The modal promotions elasticity in Table 5 is 0.233, which means 10% increase in the *frequency* of promotions yields 2.33% increase in consumer demand. Note that managers did not change the promotions depth over time or regions. Our modal promotions elasticity is much larger than 0.12 to 0.14 reported in previous studies (e.g., Kremer, Bijmolt, Leeflang, & Wieringa, 2008, Mela, Jedidi, & Bowman, 1998). Our setting represents a product in a mature product lifecycle stage, with moderate promotional support. Kremer et al. (2008) show that promotional elasticity tends to be higher in the early stage than the mature stage and in regions where products are heavily advertised. More importantly, the participating company suggests the following explanation. When they offer non-price retail promotions, consumers engage in stockpiling and purchase acceleration, which increases sales in that month but reduces it in the future months. This non-responsiveness to purchasing in the post-promotion period —"lie in wait" behavior documented by Mela et al. (1998)— in turn suppresses ad elasticity because advertising continues all year round. Hence, even non-price promotions hurt advertising efforts to build brands.

Across regions, advertising works (while promotions do not) in Andhra Pradesh, promotions work (and not advertising) in Gujarat, Tamil Nadu and Uttar Pradesh, and both work only in Karnataka and Maharashtra. In general, advertising and promotions seemingly counteract each other, inducing a negative correlation of −0.8 between ad and promotions elasticities across regions. One explanation is that most consumers in small markets (Gujarat, Tamil Nadu and Uttar Pradesh) buy unbranded commodity

Table 3
Fit and forecast.

	Andhra Pradesh	Gujarat	Karnataka	Maharashtra	Tamil Nadu	Uttar Pradesh
Retail offtake fit (MAPE[a])	5.92	16.94	15.68	10.33	16.91	18.31
Retail offtake forecast (MAPE)	9.41	18.59	13.50	9.37	14.56	21.87
Secondary sales fit (MAPE)	15.68	24.39	11.76	16.02	15.27	15.76
Secondary sales forecast (MAPE)	19.83	22.50	16.81	24.04	11.81	17.97

[a] MAPE stands for mean absolute prediction error.

Table 4
Estimation results.

Parameters	Andhra Pradesh Estimate	Andhra Pradesh t-Value	Gujarat Estimate	Gujarat t-Value	Karnataka Estimate	Karnataka t-Value	Maharashtra Estimate	Maharashtra t-Value	Tamil Nadu Estimate	Tamil Nadu t-Value	Uttar Pradesh Estimate	Uttar Pradesh t-Value
Retail offtakes noise, σ_1	**11.97**[a]	**4.91**	**8.14**	**8.72**	**11.72**	**6.29**	**18.41**	**13.48**	**6.23**	**7.86**	**7.05**	**4.70**
Secondary sales noise, σ_2	**27.60**	**6.99**	**22.46**	**6.12**	**17.05**	**5.74**	**67.64**	**5.85**	**5.14**	**5.57**	**5.21**	**2.56**
Bias, π_1	**0.11**	**7.29**	**0.20**	**7.79**	**0.10**	**4.83**	**2.58**	**10.09**	**0.02**	**5.22**	**0.12**	**5.15**
Bias, π_2	**0.12**	**6.51**	**0.27**	**6.18**	**0.11**	**4.57**	**4.38**	**10.52**	**0.02**	**4.68**	**0.13**	**5.46**
Carryover effect, λ	**0.94**	**67.51**	**0.93**	**28.62**	**0.91**	**25.50**	**0.88**	**18.54**	**0.90**	**27.49**	**0.87**	**19.89**
Synergy between distribution & advertising, γ_1	**0.06**	**8.82**	0.00	0.88	**0.03**	**2.44**	**0.001**	**4.90**	0.00	0.03	0.001	0.28
Synergy between distribution & promotion, γ_2	2.11	1.77	**3.61**	**5.09**	**3.95**	**2.77**	**0.35**	**3.68**	**4.81**	**6.89**	**3.35**	**8.15**

[a] Bold estimates are statistically significant at the 95% confidence level.

hair oil and they respond to brand's promotions. In contrast, consumers in large markets (Maharashtra, Andhra Pradesh and Karnataka) are familiar with the firm's branded hair oil and they respond to the brand's advertising.

Finally, by comparing spending from Table 2 and elasticities from Table 5, we gain insights into the company's allocation behavior. For example, advertising elasticities are negligible in Gujarat, Tamil Nadu and Uttar Pradesh, yet the company advertises heavily in Gujarat, Tamil Nadu and Uttar Pradesh. In other words, the company advertises in Gujarat, Tamil Nadu and Uttar Pradesh more than the amount they would spend based on either ad elasticity or proportionality to sales arguments. Similarly, the company advertises in Maharashtra, Andhra Pradesh, and Karnataka less than the amount they would spend based on either ad elasticity or sales proportionality. As for promotions, the company promotes *equally* frequently in all six regions, even though promotion elasticities vary from 0.036 to 0.259 across regions.

To understand this peculiarity, we note that Gujarat, Tamil Nadu and Uttar Pradesh are relatively small markets with an average volume (across all three regions) of 36.7 and 42.7 kl in offtakes and secondary sales, respectively, whereas Andhra Pradesh, Karnataka, Maharashtra are about five times larger with an average volume (across all three regions) of 124.3 and 167.5 kl in offtakes and secondary sales, respectively. In these small markets, advertising-to-sales ranges between 22.9 and 58.5 GRPs/kl; in the large markets, it ranges between 1.9 and 3.4 GRPs/kl. That is, despite ad elasticities being low, advertising intensity in small markets is about 15 times higher than that in the large markets. In contrast, promotional elasticity is larger in small markets than in the large markets. Our discussion reveals that brand managers hope that this disproportionate ad spending helps (1) grow the small markets and (2) steer consumers away from promotions by emphasizing the product benefits and branding. They do recognize the need for improving profitability by using models and data and have actively sought our guidance on the optimal budget and allocations in the presence of unreliable metrics, which we next discuss.

5. Optimal marketing mix using unreliable metrics

Given the unreliability of sales metrics, how should brand managers determine optimal advertising spending and promotional timing? How should they alter the total budget as unreliability increases? To answer such substantive issues, we formulate and solve a manager's decision-making problem taking into account the measurement noise and relative bias in the metrics.

5.1. Budgeting and allocation problem

Suppose the manager decides to spend on advertising and promotions over time as follows: $\{u_t, v_t : t \in (1,2,\ldots)\}$. This marketing plan generates a sales sequence $S(t)$, which is measured using two noisy metrics $Y_1(t)$ and $Y_2(t)$. A forward-looking manager's problem is to determine the optimal advertising strategy and promotional timing sequence so as to maximize the net present value of the stream of profits:

$$\Pi(u(t), v(t)) = \int_0^\infty e^{-\rho t}[mS(t) - u(t) - c(t)v(t)]dt, \tag{15}$$

where ρ denotes the discount rate, m is the price-cost margin per unit sold, $c(t)$ is the cost of promotion at time t. The maximization of Eq. (15) is subject to the continuous-time version of Eq. (2) given by $dS/dt = \beta_1(t)\sqrt{u(t)} + \beta_2(t)v(t) - \delta S$, where $\delta = 1 - \lambda$.

Table 5
Elasticity estimates.

	Andhra Pradesh	Gujarat	Karnataka	Maharashtra	Tamil Nadu	Uttar Pradesh
Advertising elasticity	0.254	0.001	0.105	0.016	0.001	0.001
Promotion elasticity	0.209	0.162	0.317	0.036	0.259	0.257

Since $S(t)$ is not directly observed and rather it is measured via two noisy and biased metrics, we apply Ito's lemma to obtain the stochastic evolution of the observed metrics. Applying Ito's lemma to the first row of Eq. (1), we observe that $dY_1 = \pi_1 dS + \sigma_1 dW_1$, where $W_1(t)$ is the standard Wiener process. Then, substituting $\pi_1 S(t) = Y_1(t) - \varepsilon_1(t)$ in the sales dynamics, we get $\pi_1 dS = [\pi_1\beta_1(t)\sqrt{u(t)} + \pi_1\beta_2(t)v(t) - \delta(Y_1(t) - \varepsilon_1(t))]dt = [\pi_1\beta_1(t)\sqrt{u(t)} + \pi_1\beta_2(t)v(t) - \delta Y_1]dt + \delta\sigma_1 dW_1$. Next, by substituting this expression for $\pi_1 dS$ in $dY_1 = \pi_1 dS + \sigma_1 dW_1$, we get Eq. (16), which represents the sales dynamics in the observed metric Y_1. Similarly, we derive the Eq. (17).

$$dY_1 = \left(\pi_1\beta_1(t)\sqrt{u(t)} + \pi_1\beta_2(t)v(t) - \delta Y_1\right)dt + \sigma_1(1+\delta)dW_1 \tag{16}$$

$$dY_2 = \left(\pi_2\beta_1(t)\sqrt{u(t)} + \pi_2\beta_2(t)v(t) - \delta Y_2\right)dt + \sigma_2(1+\delta)dW_2 \tag{17}$$

Thus, the measurement noise induces a stochastic control problem.

To solve this stochastic control problem, we apply the Hamilton-Jacobi-Bellman principle, which leads to a partial differential equation for the value function $V(Y_1, Y_2)$. The resulting problem is complicated because, mathematically, the optimal solution to-be-derived has to take into account four trade-offs: the present versus future (captured through the discount rate ρ), the differential effectiveness of advertising and promotions (captured through $\beta_1(t)$ and $\beta_2(t)$), the relative bias in the two metrics (captured through π), and the effects of unequal signal-to-noise ratios (captured through σ_1 and σ_2).

The goal is to derive the optimal advertising spending $u^*(t) \in (0, \infty)$ which informs how much to spend each week, and the optimal promotions indicator $v^*(t) \in \{0, 1\}$, which informs whether or not to spend on promotions given the time-varying promotional cost $c(t)$. Consequently, the control domains are mixed: continuous-valued control for advertising and binary switch for promotional timing.

5.2. Optimal advertising and promotions in the presence of unreliable metrics

Taking into account the aforementioned four tradeoffs, we solve the stochastic control problem analytically. We relegate its proof to Appendix C and present here the final results. Let $u_0^*(t)$ and $v_0^*(t)$ denote the optimal advertising and optimal promotions, respectively, in the presence of perfectly reliable metrics (i.e., with no noise or bias). Then the optimal advertising and promotions in the presence of unreliable metrics are given by

Proposition 2. $u^*(t) = u_0^*(t) \times \left(\frac{\pi_1^2\phi_1 + \pi_2^2\phi_2}{1+\pi_1^2\phi_1 + \pi_2^2\phi_2}\right)^2$ *and* $v^*(t) = \begin{cases} 1, & \text{if } v_0^*(t) \times \frac{\pi_1^2\phi_1 + \pi_2^2\phi_2}{1+\pi_1^2\phi_1 + \pi_2^2\phi_2} > c(t) \\ 0, & \text{otherwise} \end{cases}$

Proof. See Appendix C.

We refer to the expression $\frac{\pi_1^2\phi_1 + \pi_2^2\phi_2}{1+\pi_1^2\phi_1 + \pi_2^2\phi_2}$ as the correction factor (CF). It differs from the optimal weights derived in Proposition 1 and depends on the bias and metrics unreliability in a non-trivial manner. More importantly, it moderates the optimal advertising and promotions decisions under perfect reliability. To compute the correction factor, managers should apply the estimation approach in Section 3.3 to their market data, and use the estimated parameters to compute the impact and consequences of unreliability in their decision-making.

Proposition 2 fully characterizes the optimal total budget and its optimal split between advertising and promotions over time. Based on Eq. (15), we find that the total budget at any time equals $\underbrace{B(t)}_{\text{Total Budget}} = \underbrace{u(t)}_{\text{Advertising Budget}} + \underbrace{c(t)v(t)}_{\text{Promotions Budget}}$. Hence, the optimal total budget equals $B^*(t) = u^*(t) + c(t) \times v^*(t)$. Consequently, the optimal allocation ratios for advertising and promotional spending, respectively, are as follows:

$$\Lambda_1(t) = \frac{u^*(t)}{B^*(t)} \tag{18}$$

$$\Lambda_2(t) = \frac{c(t)v^*(t)}{B^*(t)} \tag{19}$$

where $(u^*(t), v^*(t))$ comes from Proposition 2.

By further analyzing the correction factor, we gain the following two insights.

Proposition 3. *As unreliability increases, first, marketing spending should be reduced. Second, this reduction is more severe for advertising than for promotions.*

Proof. CF $\frac{\pi_1^2\phi_1 + \pi_2^2\phi_2}{1+\pi_1^2\phi_1 + \pi_2^2\phi_2} < 1$. So $\frac{\partial u_t^*}{\partial \phi_i} > 0$, and $\frac{\partial v_t^*}{\partial \phi_i} > 0, i \in (1, 2)$. As measurement noise increases, the signal-to-noise ratio ϕ_i decreases and hence the optimal advertising and promotions decrease.

An insight emerging from Proposition 3 is the following. In the absence of the proposed estimation method to quantify measurement noises and in the absence of the formula for the correction factor derived in Proposition 2, brand managers have no recourse but to ignore the measurement noise. Consequently, they would act as if the metrics are perfect (i.e., noise free), which would lead to overspending on advertising and promotions as implied by Proposition 3. In other words, it pays to quantify the measurement noise, estimate signal-to-noise ratios, and then adjust the spending levels as per the correction factor. *Overconfidence in data quality is hazardous to profitability*.

Another insight emerging from Proposition 3 is the interaction effect. Specifically, it follows from Proposition 2 that the optimal advertising is proportional to the square of the correction factor, whereas the optimal promotion is linear in the correction factor. Because the correction factor is less than unity (see the proof of Proposition 3), as measurement noise increases, the reduction in advertising is faster than that required for promotions, thereby inducing the interaction effect. Consequently, unreliability *asymmetrically* alters the optimal advertising and promotions decisions.

We close this section with a graphical illustration of the above results on marketing overspending. To this end, let $\pi_1 = \pi_2 = 1$, $\phi_1 = \phi_2 = \phi = \sigma_S^2/\sigma^2$. Then we vary the signal-to-noise ratio ϕ from 1 to 10. Because metrics unreliability is the reciprocal of signal-to-noise ratio, it varies from 0.1 to 1. We compute percentage ad overspending as $100 \times (u^0 - u^*)/u^*$ and percentage promotional overspending $100 \times (v^0 - v^*)/v^*$ using Proposition 2. Fig. 2 displays how percentage overspending increases as metrics unreliability increases. This effect is *asymmetric*: advertising overspending increases at a faster rate than promotional overspending.

6. Discussion

6.1. Incorporating unreliable marketing mix metrics

We elucidate here how the proposed framework can incorporate multiple noisy marketing-mix variables. Let X_a and X_p denote the noisy metrics for advertising and promotional efforts, while and A_t and p_t correspond to their true unobserved values. Applying the errors-in-variable framework, we express $X_{at} = A_t + \epsilon_{3t}$ and $X_{pt} = p_t + \epsilon_{4t}$. Then we use these equations to augment the observation and transition equations. Specifically, the augmented observation equation to filter out the measurement noise is given by

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \\ X_{at} \\ X_{pt} \end{bmatrix} = \begin{bmatrix} \pi_1 & 0 & 0 & 0 & 0 \\ \pi_2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} S_t \\ A_t \\ p_t \\ \beta_{1t} \\ \beta_{2t} \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \\ \epsilon_{4t} \end{bmatrix}. \tag{20}$$

The corresponding transition equation is given by

$$\begin{bmatrix} S_{t+1} \\ A_{t+1} \\ p_{t+1} \\ \beta_{1,t+1} \\ \beta_{2,t+1} \end{bmatrix} = \begin{bmatrix} \lambda & \beta_{1t} & \beta_{2t} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_t \\ A_t \\ p_t \\ \beta_{1t} \\ \beta_{2t} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \gamma_1 z_t \\ \gamma_2 z_t \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \\ \eta_{3t} \\ \eta_{4t} \\ \eta_{5t} \end{bmatrix}. \tag{21}$$

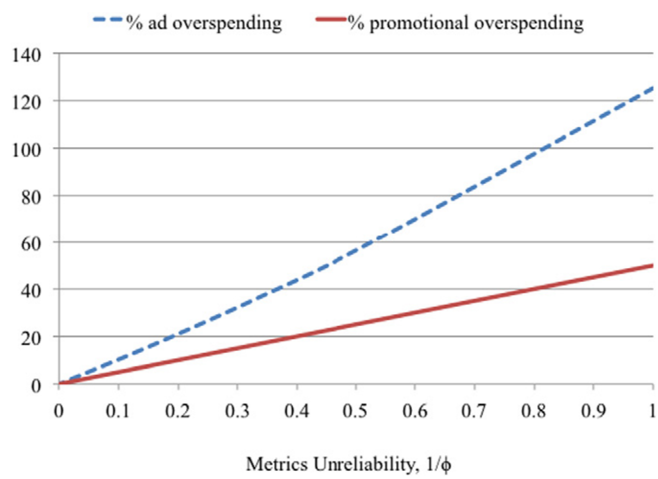

Fig. 2. Metrics unreliability and marketing overspending.

Eqs. (20) and (21) filter out the noise simultaneously from multiple inputs and multiple outputs, thus illustrating the flexibility of the proposed approach. The resulting more general dynamic system can be estimated via the extended Kalman filter (for details, see Naik, Prasad, & Sethi, 2008).

6.2. Implications for researchers

The adverse consequences of measurement noise raise serious implications. It injects inconsistency in parameter estimation – that is, no matter how large the available sample sizes is, the estimated parameters will not converge to the true parameter values (e.g., Naik & Tsai, 2000). So what should researchers do? First and foremost, they should acknowledge the possibility that metrics can be unreliable, unless proven otherwise (e.g., the estimated noise level is statistically insignificant), rather than *ex ante* (a) presuppose that "excellent data is available in my industry." Such empirically untested beliefs breed overconfidence — it does not reduce unreliability when present. Second, heed to the advice of Morrison and Silva-Risso (1995, p. G61): "…they [researchers] should explicitly consider their data as coming from the model: Observed Value = True Score + Error." Thirdly, use the methods to filter out the noise so as to restore consistency. Section 3 offers such a method to estimate dynamic models using multiple noisy and biased metrics.

How should researchers reduce unreliability in their specific applications? Currently, researchers do not combine multiple metrics even when data are available (see the last column in Table 1). More recently, Bass et al. (2007) had data on two metrics of demand (see p. 184) —minutes of call time and number of calls— but they ignored the second metric. Because a composite metric can reduce noise, we recommend using multiple independent items rather than single items. Extensive simulation studies by Diamantopoulos, Sarstedt, Fuchs, Wilczynski, and Kaiser (2012) corroborate this recommendation, and they conclude that "… opting for SI [single-item] measures in most empirical settings is a risky decision as the set of circumstances that would favor their use is unlikely to be frequently encountered in practice." Even noisier independent metrics add value; for details, see the last paragraph in Section 3.2. So don't rely solely on the usual metric if one more is available. For more details on when to use single- versus multi-item scales, we ask readers to follow the general guidelines in Diamantopoulos et al. (2012).

Another consideration is whether the importance of metrics combination wanes with technological progress? The answer is, No. To see this point, consider for illustration two independent noisy metrics $Y_1 \sim N(\mu, \sigma^2)$ and $Y_2 \sim N(\mu, \sigma^2)$. A composite metric $Y = \frac{1}{2}Y_1 + \frac{1}{2}Y_2$ has $Var(Y) = \frac{1}{4}\sigma^2 + \frac{1}{4}\sigma^2 = 0.5\sigma^2$, which is half the noise level than that in the original metrics. Now imagine that technological progress makes Y_1 half as noisy; that is, $Y_1 \sim N(\mu, \sigma^2/2)$. Then the composite metric, with weights proportional to precision (namely 2:1), $Y = \frac{2}{3}Y_1 + \frac{1}{3}Y_2$, has variance equal to $\frac{4}{9}(\sigma^2/2) + \frac{1}{9}\sigma^2 = \sigma^2/3$, which is *less than the variance of the "technologically improved" metric* Y_1 (namely $\sigma^2/2$). This example holds more generally with noise levels reduced by any factor k (instead of $k = 2$ in this example), revealing that technological improvements help, yet metrics combination itself serves as the noise reduction device. The main point of this illustration is: there exists some composite metric that is more precise than the individual metrics alone. In Proposition 1, we discover the optimal weights to construct that "some" composite metric, which depends on the extent of bias and noise in the metrics. Thus, the importance of the optimal composite metric does not wane due to advances in the technological of measurement.

6.3. Implications for managers

The managerial implications pertain to the attribution and budgeting problems. The attribution problem refers to assigning how much sales increase is due to incremental advertising or promotions, whereas the budgeting problem is to map the attribution results to obtain the appropriate spending levels on advertising and promotions. Both the attribution and budgeting problems are impacted by the unreliability of sales metrics.

In the context of attribution problems, the key question that arises is, is the observed sales increase due to advertising (or promotions) or is it due to measurement error? Measurement errors inject uncertainty that obscures the assessment of pure advertising effect, i.e., the marginal sales increase attributable to incremental advertising after controlling for noise. To hedge this uncertainty, managers should find another sales metric from a different source, even if it is noisier, to create the composite metric, which serves as the variance reduction device (see Proposition 1). Conceptually, to the extent that both noisy metrics increase (or decrease) as advertising spending varies, the less likely it is the effect of measurement noise. In other words, multiple metrics enable teasing apart the marketing-mix effectiveness from the measurement noise effect. Hence, the implication for attribution problems is that managers should *use multiple sales metrics* and *combine them appropriately* via Proposition 1.

In the context of budgeting problems, the key question that arises is, is the relation between the marketing-mix effectiveness and the optimal budget moderated by the presence of measurement noise? Proposition 3 uncovers this moderating effect of unreliability on the optimal budget. Specifically, it reveals that the optimal total budget should be reduced in the presence of unreliable metrics, and more so on advertising than on promotions. Thus, the implication for managers is simply this: *save money by reducing wasteful spending on marketing*. To accomplish this goal, managers need to estimate the measurement noise levels via the proposed approach, calculate the correction factor ($= \frac{\pi_1^2\phi_1 + \pi_2^2\phi_2}{1 + \pi_1^2\phi_1 + \pi_2^2\phi_2}$), and then adjust the marketing budget and its allocation as per Proposition 3.

6.4. Limitations and future research

Our empirical application includes the relevant marketing-mix variables, namely, advertising, promotions, and distribution activities of the hair oil brand in Indian markets. Because this brand enjoys market power, it maintains prices nearly constant over time and identical across regions. Consequently, the price effects cannot be estimated. Future researchers can study other brands that vary prices over time (or regions) by extending the proposed model appropriately. Such analysis would contribute new insights into how unreliable metrics moderate the optimal pricing decisions, for example, should managers underprice to hedge the uncertainty in latent demand?

Our empirical application illustrates that the measurement noise and bias formulated in Eq. (1) exist in the real markets. We furnish strong evidence that measurement noise and bias are large and statistically significant in all the six regions, which cover over 80% of the national sales. These results may be generalized to multiple brands and several countries. Future researchers may address new important questions. For example, managers collect sales data from multiple vendors such as GfK and Nielsen. Which vendor is more reliable? Should managers pay to obtain metrics from the less reliable vendor? The answer to the first question comes from quantifying the reliability of the two metrics. The answer to the second question requires balancing the incremental cost of purchasing the second sales metric with the corresponding savings due to the lower optimal marketing budget (via Proposition 3). We believe such studies would not only inform managers to make improved marketing decisions, but also augment our marketing understanding with new insights hitherto unavailable in the extant literature.

7. Conclusion

The adverse consequences of metrics unreliability on statistical issues (e.g., inconsistency, attenuation bias) are well known, but there exists sparse knowledge on how measurement unreliability affects strategic marketing decisions: marketing budget, allocation, or overspending? Hence we contribute three new propositions.

Proposition 1 provides a constructive way for both managers and researchers to improve reliability by combining two unreliable metrics. It reveals that the two independent noisy metrics when properly combined are better than one in terms of reliability. This result is important because researchers do not usually tend to combine metrics (see the last column of Table 1, or see Bass et al. (2007, p. 184) who had two metrics but did not combine them). It also shows that, even if technology improves the reliability of future metrics, the composite metric that combines the independent noisy metrics from yester years with the new less noisy ones has *greater* reliability than that of the “new and improved” metrics. Hence the value of Proposition 1 endures beyond technological progress.

Proposition 2 presents the optimal budget and its optimal allocation to advertising and promotions in the presence of unreliable sales metrics. It provides closed-form expression for the correction factor that managers can calculate and then adjust their budget and allocation decisions. In other words, managers can use the proposed approach to find close-to-optimum solutions by applying it in a marketplace a few times until the optimum is reached.

Proposition 3 sheds light on how overconfidence in metrics leads to overspending in marketing. Specifically, managers will overspend when they ignore measurement noise. This overspending is asymmetric— ignoring measurement noise leads to more overspending on advertising than on promotions (see Fig. 2).

Because measurement errors inject inconsistency in parameter estimation, we developed a method to control measurement errors and restore consistency in parameter estimation. Thus, managers and researchers alike should use the proposed method to obtain consistent estimates even in the presence of unreliable metrics.

Besides the above theoretical and methodological contributions, we offer new empirical results from six markets of India. The goal of our empirical application is to illustrate and confirm that the measurement noise and biases assumed in Eq. (1) do *exist in the real markets*. Empirical results furnish new evidence that the measurement noise and biases are large and statistically significant in the six regions of India (covering 80% of national sales) and in both the metrics (retail offtakes and secondary sales).

We hope managers and researchers use the proposed framework to control metrics unreliability and eliminate marketing overspending.

Appendix A. Derivation of P_t

We derive Eq. (7) by noting that $P_t = E[\omega_t \omega_t' | \Im_t]$, where

$$\begin{aligned}
\omega_t &= \alpha_t - \hat{\alpha}_t \\
&= \alpha_t - \left[\hat{\alpha}_{t|t-1} + K_t\left(Y_t - \hat{Y}_t\right)\right] \\
&= \left(\alpha_t - \hat{\alpha}_{t|t-1}\right) - K_t\left(Y_t - H\hat{\alpha}_{t|t-1}\right) \\
&= \omega_{t|t-1} - K_t\left(H\alpha_t + \epsilon_t - H\hat{\alpha}_{t|t-1}\right) \\
&= \omega_{t|t-1} - K_t H\left(\alpha_t - \hat{\alpha}_{t|t-1}\right) - K_t\epsilon_t \\
&= \omega_{t|t-1} - K_t H\omega_{t|t-1} - K_t\epsilon_t \\
&= (I - K_t H)\omega_{t|t-1} - K_t\epsilon_t
\end{aligned} \tag{A1}$$

To evaluate $P_t = E[\omega_t \omega_t' | \Im_t]$, we first multiply the cross product terms and then evaluate the expectations as follows:

$$\begin{aligned} P_t &= E\left[\omega_t \omega_t' | \Im_t\right] \\ &= E\left[\left\{(I - K_t H)\omega_{t|t-1} - K_t \epsilon_t\right\}\left\{(I - K_t H)\omega_{t|t-1} - K_t \epsilon_t\right\}'\right] \\ &= (I - K_t H) E[\omega_{t|t-1}\omega_{t|t-1}'(I - K_t H)' - (I - K_t H) E\left[\omega_{t|t-1}\epsilon_t'\right] K_t' - K_t E\left[\epsilon_t \omega_{t|t-1}'(I - K_t H)' + K_t E \epsilon_t \epsilon_t'\right] K_t' \\ &= (I - K_t H) P_{t|t-1}(I - K_t H)' + K_t R K_t' \end{aligned} \tag{A2}$$

where the last equality follows because the middle terms vanish given the independence across periods.

Appendix B. Proof of Proposition 1

We derive the expressions for k_1^* and k_2^* to combine the unreliable metrics Y_1 and Y_2. Recall that, abstracting from the dynamics to suppress time subscripts (but not ignoring it in the estimation Section 3.3), $\hat{S} = k_1^* Y_1 + k_2^* Y_2$, where $\begin{bmatrix} k_1^* & k_2^* \end{bmatrix}$ are the elements from the first row of the matrix $K^* = PH'(HPH' + R)^{-1}$.

Let $\{p_{ij}\}$ denote the elements of the 3×3 matrix P, with $p_{11} = \sigma_S^2$, the 2×3 matrix $H = \begin{bmatrix} \pi_1 & 0 & 0 \\ \pi_2 & 0 & 0 \end{bmatrix}$, and the 2×2 matrix $R = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$. Then the 2×2 matrix $(HPH' + R)$ equals $\begin{bmatrix} \pi_1^2\sigma_S^2 + \sigma_1^2 & \pi_1\pi_2\sigma_S^2 \\ \pi_1\pi_2\sigma_S^2 & \pi_2^2\sigma_S^2 + \sigma_2^2 \end{bmatrix}$, and so its inverse is $\begin{bmatrix} \pi_2^2\sigma_S^2 + \sigma_2^2 & -\pi_1\pi_2\sigma_S^2 \\ -\pi_1\pi_2\sigma_S^2 & \pi_1^2\sigma_S^2 + \sigma_1^2 \end{bmatrix}/D$, where $D = \pi_2^2\sigma_S^2\sigma_1^2 + \pi_1^2\sigma_S^2\sigma_2^2 + \sigma_1^2\sigma_2^2$. Also the matrix $PH' = \begin{bmatrix} \pi_1\sigma_S^2 & \pi_2\sigma_S^2 \\ \pi_1 p_{21} & \pi_2 p_{21} \\ \pi_1 p_{31} & \pi_2 p_{31} \end{bmatrix}$.

Consequently, we obtain

$$k_1^* = \frac{\pi_1\phi_1}{1 + \pi_1^2\phi_1 + \pi_2^2\phi_2} \tag{B1}$$

where $\phi_i = \sigma_S^2/\sigma_i^2$ represents the signal-to-noise ratio in the metric Y_i. Similarly, we obtain

$$k_2^* = \frac{\pi_2\phi_2}{1 + \pi_1^2\phi_1 + \pi_2^2\phi_2} \tag{B2}$$

The closed-form expressions in (B1) and (B2) furnish the optimal weights to combine the unreliable metrics, thus proving Proposition 1.

Appendix C. Optimal budget and allocation in the presence of unreliable metrics

We solve the marketing budgeting and allocation problem stated in Eq. (15):

$$Max_{(u,v)}\left\{\Pi(u(t), v(t)) = \int_0^\infty e^{-\rho t}[mS(t) - u(t) - c(t)v(t)]dt\right\}$$

subject to

(i) $\frac{dS}{dt} = \beta_1(t)\sqrt{u(t)} + \beta_2(t)v(t) - \delta S$,

(ii) $\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \pi_1 S_t \\ \pi_2 S_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$.

The presence of measurement noise in the metrics introduces uncertainty, which is represented by the Wiener processes in Eqs. (16) and (17). Consequently, to maximize the total profit $\Pi(u,v)$, we need to solve a stochastic control problem.

To this end, we formulate the Hamilton-Jacobi-Bellman equation as follows:

$$\rho V = Max\left[(m(k_1^*Y_1 + k_2^*Y_2) - u - cv) + \dot{V}_1(\pi_1\beta_1\sqrt{u} + \pi_1\beta_2 v - \delta Y_1) + \dot{V}_2(\pi_2\beta_1\sqrt{u} + \pi_2\beta_2 v - \delta Y_2) + 0.5\sigma_1^2(1+\delta)^2\ddot{V}_1 + 0.5\sigma_2^2(1+\delta)^2\ddot{V}_2\right], \tag{C1}$$

where we suppress the time argument for clarity, use the result in Proposition 1, and denote the value function by $V(Y_1, Y_2)$, with its first partial derivatives as $\dot{V}_i = \partial V/\partial Y_i$ and the second partial derivatives as $\ddot{V}_i = \partial^2 V/\partial Y_i^2$ for each metric $i\in(1,2)$. Thus Eq. (C1) is a second-order partial differential equation.

Next, to determine the optimal advertising, we differentiate the right hand side of Eq. (C1) with respect to u and get the first-order condition (FOC) as follows:

$$-1+\frac{\dot{V}_1\pi_1\beta_1}{2\sqrt{u}}+\frac{\dot{V}_2\pi_2\beta_1}{2\sqrt{u}}=0, \tag{C2}$$

which upon re-arrangement gives the optimal advertising:

$$u^*(t)=\left(0.5\beta_1\left(\pi_1\dot{V}_1+\pi_2\dot{V}_2\right)\right)^2. \tag{C3}$$

Based on previous research (e.g., Aravindakshan et al., 2012), we specify the value function $V(Y_1,Y_2)=v_0+v_1Y_1+v_2Y_2$. Consequently, $\dot{V}_1=v_1, \dot{V}_2=v_2,$ and $\ddot{V}_i=0.$ To further express (v_1,v_2) in terms of the model parameters, we replace $u^*=(0.5\beta_1(\pi_1v_1+\pi_2v_2))^2$ in the HJB Eq. (C1) and equate the coefficients for (Y_1,Y_2) on both sides of the equality. Simplifying the resulting algebra, we obtain.

$$v_1=\frac{mk_1^*}{\rho+\delta}, \text{and } v_2=\frac{mk_2^*}{\rho+\delta}. \tag{C4}$$

Using the expressions in (C4) and the optimal weights (k_1^*,k_2^*) from Proposition 1, we thus characterize the optimal advertising strategy in the presence of unreliable metrics:

$$u^*(t)=u_0^*(t)\times\left(\frac{\pi_1^2\phi_1+\pi_2^2\phi_2}{1+\pi_1^2\phi_1+\pi_2^2\phi_2}\right)^2, \tag{C5}$$

where $u_0^*(t)=(0.5m\beta_1(t)/(\rho+\delta))^2$.

Finally, to determine the optimal promotion timings, we differentiate the right hand side of Eq. (C1) with respect to v to get

$$-c+\beta_2\left(\pi_1\dot{V}_1+\pi_2\dot{V}_2\right), \tag{C6}$$

which is not a function of the decision variable, $v(t)$.Hence the optimal solution belongs to the class of bang-bang controls, indicating when to switch "on" or "off" based on the switching function specified by Eq. (C6). Using the expressions in (C4) and the optimal weights (k_1^*,k_2^*) from Proposition 1, we thus characterize the optimal promotion strategy in the presence of unreliable metrics:

$$v^*(t)=\begin{cases}1, \text{ if } v_0^*(t)\times\dfrac{\pi_1^2\phi_1+\pi_2^2\phi_2}{1+\pi_1^2\phi_1+\pi_2^2\phi_2}>c(t)\\ 0, \text{otherwise}\end{cases} \tag{C7}$$

where $v_0^*(t)=m\beta_2(t)/(\rho+\delta)$.

The closed-form expressions in (C5) and (C7) comprise the optimal marketing allocations in the presence of unreliable data, thus proving Proposition 2.

Appendix D. Identification

We discuss the identification of parameters in Eq. (1). In static structural equation models (e.g., LISREL), Y_i is mean-centered, $E[S]=0$, and $\sigma_S^2=1$. Then, the sample covariance matrix $Cov(Y_1,Y_2)$ yields the three sample moments (i.e., two variances and one covariance), which offer three equations to solve for the five unknown parameters $(\pi_1,\pi_2,\sigma_1^2,\sigma_2^2,\rho)'$. Specifically, the three equations are as follows:

$$\begin{aligned}&Var(Y_1)=\pi_1^2+\sigma_1^2\\&Var(Y_2)=\pi_2^2+\sigma_2^2\\&Cov(Y_1,Y_2)=\rho\sigma_1\sigma_2+\pi_1\,\pi_2.\end{aligned}$$

This shortage of information — fewer equations than unknowns — is the source of non-identification in static models. Hence, two unknown parameters must be fixed, leading to the usual choices of $\pi_1=1$ and $\rho=0$ so that $(\pi_2,\sigma_1^2,\sigma_2^2)'$ can be identified.

In contrast, all five parameters $(\pi_1,\pi_2,\sigma_1^2,\sigma_2^2,\rho)'$ are identified in dynamic models. To see this point, we recall that the dynamic equations such as those in Eqs. (2) or (3) can be expressed in the canonical vector-matrix transition equation $\alpha_t=T_t\alpha_{t-1}+c_t+\nu_t$ in the state space modeling framework (see Harvey, 1994). Then the state vector α_t is a random variable, whose *first two moments are identified by the Kalman filter* via the celebrated closed-form recursions of the mean state vector a_t and the covariance matrix

P_t (e.g., see Naik et al., 1998, Appendix B). Because the first equation in Eqs. (2) or (3) pertains to latent sales, the first element of the mean state vector is $a_1 = E[S]$ and the first diagonal element of the covariance matrix $P_{11} = \sigma_S^2$. Because $E[S] = a_1 \neq 0$, we get the five equations as follows:

$$\begin{aligned}
&E(Y_1) = \pi_1 E(S) \\
&E(Y_2) = \pi_2 E(S) \\
&Var(Y_1) = \pi_1^2 \sigma_S^2 + \sigma_1^2 \\
&Var(Y_2) = \pi_2^2 \sigma_S^2 + \sigma_2^2 \\
&Cov(Y_1, Y_2) = \rho\sigma_1\sigma_2 + \pi_1\pi_2\sigma_S^2
\end{aligned}$$

Consequently, *the first two equations identify* $\pi_i = \overline{Y}_i/a_1$, for $i = 1,2$,and $\overline{Y}_i = E[Y_i]$ are the ensemble averages. In other words, $E[Y_i]$ is an average over the ensemble of time paths from the stochastic process Y_{it} for every t; it is not an average over a time span and the moments are available for every t. Because we know σ_S^2 from the first element of P_t and have identified (π_1,π_2) above, the remaining three equations can be solved for the three unknown parameters $(\sigma_1^2,\sigma_2^2,\rho)$. Thus, all five unknown parameters are identified in dynamic models with bias and noise in each metric. The intuition for this result is that repeated observations over time permit identification because the Kalman filter recursions furnish the moments of a state vector, which are not available in static models.

References

Abraham, M. M., & Lodish, L. M. (1993). An implemented system for improving promotion productivity using store scanner data. *Marketing Science*, *12*(3), 248–269.
Adcock, R. J. (1877). Note on the method of least squares. *The Analyst*, *4*(6), 183–184.
Adcock, R. J. (1878). A problem in least squares. *The Analyst*, *5*(2), 53–54.
Aravindakshan, A., Peters, K., & Naik, P. A. (2012). Spatiotemporal allocation of advertising budgets. *Journal of Marketing Research*, *49*(1), 1–14.
Assael, H. (1967). Comparison of brand share data by three reporting systems. *Journal of Marketing Research*, *4*(4), 400–401.
Ataman, M. B., Van Heerde, H. J., & Mela, C. F. (2010). The long-term effect of marketing strategy on brand sales. *Journal of Marketing Research*, *47*(5), 866–882.
Balakrishnan, R. (2011). *Nielsen audit back with refinements to address retail grievances. Enough to reassure skeptics?* Economic Times (Sept 21st 2011).
Bass, F. M., Bruce, N., Majumdar, S., & Murthi, B. P. S. (2007). Wearout effects of different advertising themes: A dynamic Bayesian model of the advertising-sales relationship. *Marketing Science*, *26*(2), 179–195.
Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Operations Research*, *20*(4), 451–468.
Beegle, K., Weerdt, J., Friedman, J., & Gibson, J. (2010). *Methods of household consumption measurements through surveys: Experimental results from Tanzania. World Bank Developmental Research Group, Policy Research Working Paper, Vol. 5501*.
Blattberg, R. C., & Levin, A. (1987). Modeling the effectiveness and profitability of trade promotions. *Marketing Science*, *6*(2), 124–146.
Bronnenberg, B. J., & Mahajan, V. (2001). Unobserved retailer behavior in multimarket data: Joint spatial dependence in market shares and promotion variables. *Marketing Science*, *20*(3), 284–299.
Brown, R. G. (1973). A model for measuring the influence of promotion on inventory and consumer demand. *Journal of Marketing Research*, *10*(4), 380–389.
Bruce, N. I., Peters, K., & Naik, P. A. (2012). Discovering how advertising grows sales and builds brands. *Journal of Marketing Research*, *49*(6), 793–806.
Cai, Z., Naik, P. A., & Tsai, C. -L. (2000). Denoised least squares estimators. *Statistica Sinica*, *10*(4), 1231–1242.
Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434–449.
Grace, J. B., & Bollen, K. A. (2008). Representing general theoretical concepts in structural equation models: The role of composite variables. *Environmental and Ecological Statistics*, *15*, 191.
Grewal, R., Cote, J. A., & Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing Science*, *23*(4), 519–529.
Hanssens, D., Parsons, L., & Schultz, R. (2001). *Market response models: Econometric and time series analysis* (2nd ed). Boston, Mass.: Kluwer Academic Publishers.
Hanssens, D. M. (1998). Order forecasts, retail sales, and the marketing mix for consumer durables. *Journal of Forecasting*, *17*, 327–346.
Harvey, A. C. (1994). *Forecasting, structural time series models and the Kalman filter*. New York, N.Y: Cambridge University Press.
Jöreskog, K. G. (1970). *A general method for estimating a linear structural equation system. ETS Research Report Series, Vol. 2*.
Kappe, E., & Stremerch, S. (2016). Drug detailing and doctors' prescription decisions: The role of information content in the face of competitive entry. *Marketing Science*, *35*(6), 915–933.
Kolsarici, C., & Vakratsas, D. (2010). Category versus brand-level advertising messages in a highly regulated environment. *Journal of Marketing Research*, *47*(6), 1078–1089.
Kremer, S. T. M., Bijmolt, T. H. A., Leeflang, P. S. H., & Wieringa, J. E. (2008). Generalizations on the effectiveness of pharmaceutical promotional expenditures. *International Journal of Research in Marketing*, *25*(4), 234–246.
Kuehn, A. A., McGuire, T. W., & Weiss, D. L. (1966). *Measuring the effectiveness of advertising. Proceedings of the Fall Conference, American Marketing Association, Vol. 185*.
Leeflang, P. S. H., & Olivier, A. J. (1985). Bias in consumer panel and store audit data. *International Journal of Research in Marketing*, *2*(1), 27–41.
Leone, R. P. (1995). Generalizing what is known about temporal aggregation and advertising carryover. *Marketing Science*, *14*(3), G141–G150.
Lindley, D. V. (1947). Regression lines and the linear functional relationship. *Journal of the Royal Statistical Society*, *9*(2), 218–244.
Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91.
Mela, C. F., Jedidi, K., & Bowman, D. (1998). The long-term impact of promotions on consumer stockpiling behavior. *Journal of Marketing Research*, *35*(2), 250–262.
Morgan, B. J. T. (2005). Antithetic variable. In P. Armitage, & T. Colton (Eds.), *Encyclopedia of biostatistics (Second Edition)Vol. 2*. Wiley Online.
Morrison, D. G., & Silva-Risso, J. (1995). A latent look at empirical generalizations. *Marketing Science*, *14*(3), G61–G70 Part 2 of 2.
Naik, P. A., Mantrala, M. K., & Sawyer, A. G. (1998). Planning media schedules in the presence of dynamic advertising quality. *Marketing Science*, *17*(3), 214–235.
Naik, P. A., Prasad, A., & Sethi, S. P. (2008). Building brand awareness in dynamic oligopoly markets. *Management Science*, *54*(1), 129–138.
Naik, P. A., & Raman, K. (2003). Understanding the impact of synergy in multimedia communications. *Journal of Marketing Research*, *40*(4), 375–388.
Naik, P. A., & Tsai, C. -L. (2000). Controlling measurement errors in models of advertising competition. *Journal of Marketing Research*, *37*(1), 113–124.
Narayanan, S., Desiraju, R., & Chintagunta, P. K. (2004). Return on investment implications for pharmaceutical promotional expenditures: The role of marketing-mix interactions. *Journal of Marketing*, *68*(3), 90–105.
Neslin, S. A., & Shoemaker, R. W. (1983). A model for evaluating the profitability of coupon promotions. *Marketing Science*, *2*(4), 361–388.
Pauwels, K., Erguncu, S., & Yildirim, G. (2013). Winning hearts, minds, and sales: How marketing communications enters the purchase process in emerging and mature markets. *International Journal of Research in Marketing*, *30*(1), 57–68.
Pearson, K. (1901). On lines and planes of closest to Systems of Points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572.
Ramani, G., & Kumar, V. (2008). Interaction orientation and firm performance. *Journal of Marketing*, *72*(1), 27–45.

Roshwalb, I. (1970). Two methods of estimating sales from a sample of retail outlets. *Journal of Marketing Research*, *7*(3), 396–398.
Sethuraman, R., Tellis, G. J., & Briesch, R. A. (2011). How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research*, *48*(3), 457–471.
Shoemaker, R., & Pringle, L. G. (1980). Possible biases in parameter estimation with store audit data. *Journal of Marketing Research*, *17*(1), 91–96.
Simon, J. L., & Arndt, J. (1980). The shape of the advertising function. *Journal of Advertising Research*, *20*(3), 11–28.
Srinivasan, S., & Hanssens, D. M. (2009). Marketing and firm value: Metrics, methods, findings, and future directions. *Journal of Marketing Research*, *46*(3), 293–312.
Srinivasan, S., Vanhuele, M., & Pauwels, K. (2010). Mind-set metrics in market response models: An integrative approach. *Journal of Marketing Research*, *47*(4), 672–684.
Stanton, J. L., & Tucci, L. A. (1982). The measurement of consumption: A comparison of surveys and diaries. *Journal of Marketing Research*, *19*(2), 274–277.
Steenkamp, J. -B. E. M., & de Jong, M. G. (2010). A global investigation into the constellation of consumer attitudes toward global and local products. *Journal of Marketing*, *74*(6), 18–40.
Wedel, M., & Pieters, R. (2008). Eye tracking for visual marketing. *Foundations and Trends in Marketing*, *1*(4), 231–320.
White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, *50*, 1–25.
Wind, Y., & Lerner, D. (1979). On the measurement of purchase data: Surveys versus purchase diaries. *Journal of Marketing Research*, *16*(1), 39–47.