DYNAMIC MARKETING MIX ALLOCATIONS IN THE PRESENCE OF UNRELIABLE DATA FROM INDIAN MARKETS

Shrihari Sridhar
Assistant Professor of Marketing
Smeal College of Business
Pennsylvania State University
E-mail: sus55@psu.edu

Prasad A. Naik
Professor of Marketing
Graduate School of Management
University of California Davis
Email: panaik@ucdavis.edu

Ajay Kelkar
Co-Founder and Chief Operating Officer
Hansa Cequity Pvt. Ltd.
Mumbai 400070
Email: ajay.kelkar@cequitysolutions.com

DYNAMIC MARKETING MIX ALLOCATIONS IN THE PRESENCE OF UNRELIABLE DATA FROM INDIAN MARKETS

Abstract

Mom and pop stores, who lack bar code scanning technology to track brand sales, comprise 93% of the Indian retail universe. Consequently, brand managers measure consumer demand by using a survey-based metric called *retail offtakes*, which fallibly indicates the quantity sold by retailers. They complement it with an internal metric called *secondary sales*, which fallibly indicates the quantity bought by retailers. Our analysis shows that the direct application of standard marketing-mix models to these noisy metrics individually or its convex combinations lead to inaccurate estimates of advertising and promotion elasticities. Thus we encounter the questions: How to recover correct parameter estimates using unreliable data? How to optimally combine multiple metrics to infer consumer demand? How to optimally allocate marketing investments in the presence of unreliable data? To address these issues, we formulate a new marketing-mix model, which denoises and combines the fallible metrics optimally to assess the effectiveness of marketing-mix activities. Our simulation results illustrate that the standard approach fails, whereas the proposed approach recovers the marketing-mix effectiveness accurately. Next, we analyze the market data from six political regions in India to furnish the first empirical evidence on data unreliability via the presence of bias and measurement noise in multiple sales metrics. Also, we discover the existence of distribution synergies with both advertising and promotion. Moreover, we derive closed-form analytical expressions for optimal advertising and promotion investments in the presence of unreliable metrics. Finally, we demonstrate that overconfidence due to the presumption that the metrics are reliable results in over-spending on advertising and promotion.

Keywords: Emerging Markets; Kirana Stores; Stochastic Control Theory; Bang-Bang Control; Measurement Noise; Marketing-Mix Allocations.

1. Introduction

Indian consumers' spending will be the 5th largest by 2025, surpassing Germany and crossing 8 trillion US dollars at purchasing parity, according to the McKinsey Global Institute (2007). This study predicates the growth on two forces: household disposable incomes rise at 5.8% annually, and the "middle-class" expands from the present 5% to 40% of the country's billion plus population. This emerging opportunity attracts multinational retailers with single brands (e.g., IKEA, Apple) and multiple brands (e.g., Wal-Mart, Tesco). Their entry brings with them new direct investments, business processes, and marketing know-how. An important know-how for marketing managers is to be able to measure consumer demand, assess the impact of various factors influencing it, and forecast it over time and across regions. For example, companies can use choice models to analyze household purchase behavior or marketing-mix models to analyze aggregate market sales, thereby assessing the effectiveness of advertising, promotion, or distribution and then allocating resources optimally to various activities, over time, and across regions. While this know-how is invaluable, Don Schultz (2012, p. 13) cautions marketing managers,

> "*If you want to market in another region or country, first learn what's available there … importing marketing systems from the West simply doesn't work in emerging markets*."

The main challenge to successfully importing extant models is the ability to tackle unreliability of data. To gain appreciation for data unreliability, we briefly describe how retail markets operate in India. Indian consumers buy from *kirana* shops, which are hole-in-the-wall retail stores, over 95% of these 14 million shops are smaller than 500 square feet, run by family members, sell a limited assortment of goods in small quantities, offer much-needed credit and home-delivery services to consumers, and don't use information technology to track inventory or sales.[1] This retail sector is referred to as "unorganized" and it accounts for over 93% of the country's $470 billion industry (Schumpeter 2011), and the rest is attributed to organized sector of supermarkets as in the developed countries. This fragmented universe of 14 million shops lack bar code scanning technology to track brand sales. Consequently, marketing managers commission market research companies to sample several shops nationwide and obtain the projected monthly retail sales, which they call *retail offtakes*— it serves as an error-prone proxy for consumer demand for brands.

The second challenge is to reconcile multiple metrics for brand sales. To mitigate unreliability in the retail offtakes data, managers also use internal sales-force reports to determine *secondary sales*, which refers to the quantity bought by the retailers. Managers feel that secondary sales data may be more reliable because it comes from the "internal" sources. Yet it is likely biased because it is not close to consumers' decision stage, unlike retail offtakes. For example, secondary sales data could also reflect the role played by sales contests in influencing quantity bought by retailers. Moreover, the two metrics don't match on a monthly basis, although each reflects the common latent demand fallibly. Therefore, managers need a way to reconcile unreliable metrics of brand sales when assessing the impact of marketing mix activities.

Because these challenges are inherent to Indian markets, no study in the extant marketing science literature addresses them (e.g., see Hanssens, Parsons and Schultz 2001). Hence we know neither the extent or nature of data unreliability, nor the consequences of ignoring measurement noise when estimating advertising or promotion effectiveness. More importantly, if

[1] For more details, see http://en.wikipedia.org/wiki/Retailing_in_India#cite_ref-IndianRealtyNews_9-0. Accessed June 28, 2012.

estimated effectiveness is inaccurate, how do we recover the correct estimates using biased and unreliable metrics? How do we optimally combine multiple noisy metrics to infer the true demand? Finally, how do we determine the optimal spending on advertising and promotion by accounting for multiple unreliable metrics?

To address these issues, we formulate a new marketing-mix model that incorporates the role of multiple unreliable metrics. In addition, we illustrate that the standard marketing-mix models don't work in the presence of unreliable data from Indian markets, as Don Schultz (2012) foretold. Even if we create composite sales by combining the two metrics, the extant model fails to recover marketing-mix elasticities. Then, we show that the proposed approach to filter out measurement noise from multiple metrics recovers the marketing mix effects accurately. Next, we solve the resulting stochastic control problem to derive closed-form analytical expressions for optimal advertising spending and promotion timing. Finally, we deduce propositions that shed light on how bias and noise in metrics moderate the optimal decisions with reliable data. Specifically, we prove that overconfidence in the metrics —the presumption that the metrics are reliable— enhances spending on advertising and promotion. Hence we derive a *correction factor*, which depends on the extent of bias and noise in metrics. In practice, managers can calculate this correction factor using the estimated parameters and adjust their marketing spending accordingly.

The rest of the paper is organized as follows. Section 2 presents a motivating example with real data from Indian markets to illustrate that the standard marketing-mix model does not work. Section 3 formulates the new model, and Section 4 conducts simulations to show the proposed approach recovers the true parameters. Section 5 presents the empirical analysis, and

Section 6 derives new propositions on the effects of data unreliability on optimal marketing allocations. Section 7 concludes the paper by summarizing the key takeaways.

2. Motivating Example

We seek to understand the consequences of "importing" the standard marketing-mix model to Indian markets. We relegate the data description to Section 5 and present here the results on the estimated elasticities from the standard model:

$$S_t = \lambda S_{t-1} + \beta_1 \sqrt{u_t} + \beta_2 v_t + \eta_t, \tag{1}$$

where S_t is consumer demand measured by either monthly secondary sales or retail offtakes, λ is the carryover effect, (β_1, β_2) are advertising and promotion effectiveness, (u_t, v_t) are advertising spending and a promotion "on-off" indicator, respectively, and η_t denotes the normal errors in demand specification. The squared root captures the diminishing returns to advertising (see Simon and Arndt 1980), which means the incremental sales from additional advertising diminish as spending levels increase.

What would be the elasticities if we apply regression to the unreliable data on secondary sales or retail offtakes as the *dependent* variable? Recall that the theory of errors-in-variables predicts that the estimated parameters are unbiased despite measurement noise in the dependent variables (e.g., see Greene 1993, Ch. 9). Our results differ from this claim, thus identifying a boundary condition when this theory does not hold. To this end, we estimated the standard regression model with retail offtakes (or secondary sales) as the dependent variable. To make meaningful comparisons later with the to-be-proposed model, we also include the role of distribution synergies with advertising and promotion. Specifically, we let distribution enhance the effectiveness of advertising and promotion; i.e., $\beta_{i,t} = \gamma_i z_{t-1}$, where z_t is the number of stores where this brand is available.

[Insert Table 1 about here]

Table 1 presents the estimated elasticities for advertising and promotion in the two political regions of India: Karnataka and Maharashtra. For Karnataka, advertising elasticity is 0.005 based on the secondary sales metric and 0.017 from the retail offtakes metrics, revealing that they differ by a factor of 3.4. Similarly, promotion elasticity differs by a factor of 15.5 based on whether we use secondary sales or retail offtakes. Even the relative effectiveness via the ratios of promotion to advertising elasticities differs widely. Specifically, promotion is twice as effective as advertising based on retail offtakes, whereas it is 108 times as effective as advertising when secondary sales is the dependent variable. Such a large variation creates a dilemma for managers on which metric to use to assess marketing effectiveness for allocating budgets to marketing activities.

For Maharashtra, advertising elasticity of 0.020 or 0.024 represents a reasonable variation, but the promotion elasticity differs by a factor of 51.5 based on which metric is used. Relative effectiveness also varies widely: promotion is about half as effective as advertising based on retail offtakes, but it is 25.8 times as effective based on secondary sales. Such a large variation reinforces the dilemma that managers face on the metrics to use, and so they may consider creating a composite metric by averaging the two sales metrics.

To understand the efficacy of this alternative approach, we create a *composite sales* metric $\bar{S}_t = w \times$ Retail Offtakes + $(1 - w) \times$ Secondary Sales. When the weight $w = 0.5$ we average the two metrics; for other weights we obtain asymmetric convex combinations. Then we repeat the above analysis, present the results in the lower panel of Table 1, and observe a pattern of conflicting results as described above. Specifically, advertising elasticity is negative for Karnataka, which is neither correct nor interpretable. It is positive and stable for Maharashtra, suggesting any weight would work. However, the same conclusions do not hold for promotion elasticity, which varies from 0.112 to 0.580 depending on the weights used.

In sum, the motivating example reveals that (i) elasticity estimates vary dramatically based on which metric is used, (ii) relative effectiveness of advertising to promotion differs too much to make reliable budget allocation decisions, and (iii) averaging the metrics does not resolve the dual issues of filtering noise and combining information. Thus, the standard marketing-mix model does not work when applied to sales data from Indian markets. As Don Schultz (2012) suggests, for successfully importing extant models, we need to adapt them to the prevailing market conditions, which we next address.

3. Model Development

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

3.1 Controlling Data Unreliability

As described in the Introduction, metrics are unreliable due to the absence of bar code scanning technology. Unreliability comprises of the two aspects: measurement noise in each metric and the relative bias across multiple metrics. We model the two aspects of unreliability using an errors-in-variables framework to ascertain whether noise and bias are statistically significant or not. Indeed, ignoring unreliability *ex ante* is not innocuous as the motivating example illustrates.

Let Y_{1t} and Y_{2t} denote the observed retail offtakes and secondary sales at time *t* from a given region (e.g., Karnataka), and ε_{jt} (*j* = 1, 2) be the measurement errors in each metric. Because both the metrics reflect common underlying consumer demand, S_t, we model the errors in the metrics as follows:

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

where π measures the bias in secondary sales relative to the retail offtakes, which is normalized to unity because of its proximity to consumers' purchasing stage. The error vector $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ follows a bivariate normal with zero means and the error variances σ_j^2 (*j* = 1, 2) are arranged diagonally in a matrix R. Equation (2) thus formally incorporates the notion of noisy metrics, whose presence can be ascertained via the significance of variances.

Marketing-mix activities such as advertising and promotion drive the latent consumer demand via Equation (1). In addition, it includes the carryover effect λ from the lagged demand, which the previous research (e.g., Leone 1995, Hanssens et al. 2001) indicates to be important due to the inter-temporal influence of marketing actions (e.g., past advertising or promotion effects).

In standard marketing-mix models, the effectiveness of advertising and promotion 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 (Naik and 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. Because the exact nature of time variation is not known, recent studies specify random walk evolution (e.g., Kolsarici and Vakratsas 2010). That is, $\beta_t = \beta_{t-1} + \eta_t$, which parsimoniously captures non-monotonic dynamics. Second, recent studies suggest that other marketing activities may enhance the effectiveness of advertising, a phenomenon known as synergy (e.g., Naik and Raman 2003, Narayanan, Desiraju, and Chintagunta 2004). We extend this notion to capture synergies due to distribution intensity, which is hitherto less explored in the extant literature. In other words, if a brand is more widely available, then its advertising and promotion activities are likely more effective (i.e., $\beta_{i,t+1} = \gamma_i z_t$).

We incorporate both the extensions in Equation (1) 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 a trivariate normal $N(0, Q)$ with zero means and the covariance matrix Q. The variance $\sigma^2_{\eta_1}$ captures the unexplained portion of the variation in the true demand S_t.

Equation (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 Equation (3) is called the transition matrix T_t. We link the state vector to the unreliable metrics $Y_t = (Y_{1t}, Y_{2t})'$ in Equation (1) via the observation equation:

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

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

3.2 Combining Multiple Metrics Optimally

As noted in the Introduction, the two metrics do not match exactly; that is, $Y_{1t} \neq Y_{2t}$. Hence we need to combine the information in the unreliable metrics to estimate the consumer demand $\hat{S}_t$ (which is the first element of the state vector $\hat{\alpha}_t$). One way to combine multiple metrics is to update the estimates proportional to the forecasting errors as follows:

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

where the 3 x 2 time-varying matrix K_t is to be determined. In other words, the elements $\{k_{11,t}, k_{12,t}\}$ in K_t are the weights placed on the forecasting errors to obtain $\hat{S}_t$. Similarly, using the other elements in K_t, Equation (5) updates the estimates of advertising and promotion effectiveness.

We seek to determine 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, if $\nu_t = \alpha_t - \hat{\alpha}_t$, then $E(\nu_t) = 0$ yields the unbiased estimates. Furthermore, their mean squared error is given by,

$$\begin{aligned} J_t &= E[(S_t - \hat{S}_t)^2 + (\beta_{1,t+1} - \hat{\beta}_{1,t+1})^2 + (\beta_{2,t+1} - \hat{\beta}_{2,t+1})^2] \\ &= E[\nu_t'\nu_t] = E[Tr(\nu_t'\nu_t)] = Tr(E[\nu_t\nu_t']) \\ &= Tr(P_t), \end{aligned} \tag{6}$$

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

$$P_t = (I - K_t H)P_{t-1}(I - K_t H)' + K_t R K'_t. \tag{7}$$

Finally, to bring the estimates closest to the true values, we choose the matrix K_t that minimizes Equation (6). Recalling that $\partial Tr(ABA') = 2AB\partial A$ for symmetric B, we obtain the first order condition:

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

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

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

The last equality in Equation (9) provides the optimal weights to combine the multiple metrics Y_t in Equation (5). By suppressing the time subscripts for clarity, we furnish closed-form expressions for the optimal weights 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 $k_1^* = \frac{\phi_1 + \pi(1-\pi^2)\phi_1\phi_2}{1+\pi(\phi_1+\phi_2)}$ and $k_2^* = \frac{\pi\phi_2 + \pi(\pi^2-1)\phi_1\phi_2}{1+\pi(\phi_1+\phi_2)}$ 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 is the variance of consumer demand.

Proof. *See Appendix B.*

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

Second, the optimal weights do not sum to unity. Hence the composite sales constructed via any convex combination of the metrics lead to an incorrect estimate of consumer demand, which explains the inaccurate results in the motivating example.

Finally, and most importantly, the derivation of optimal weights does not require the assumption that the measurement errors are normally distributed. In other words, the weights given in the proposition are optimal across any distribution of measurement errors with finite moments.

3.3 Parameter Estimation and Robust Inference

To assess consumer demand and marketing-mix effectiveness, we apply Equations (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,

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

where $p(\cdot \mid \cdot)$ 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 Equation (4), we find the conditional mean $\hat{Y}_t = E[Y_t \mid \Im_{t-1}] = Ha_{t|t-1}$, so the innovation errors $(Y_t - \hat{Y}_t)$ are distributed with zero mean and the covariance matrix $F_t = HB_{t|t-1}H' + R$, where $(a_{t|t-1}, B_{t|t-1})$ are the conditional means and covariances of the "prior" state vector $\alpha_t|\Im_{t-1}$. We obtain its moments via Equations (3) and (4). Specifically, $a_{t|t-1} = T_t a_{t-1} + d_t$ and $B_{t|t-1} = T_t B_{t-1} T_t' + Q$, where (a_{t-1}, B_{t-1}) are the conditional means and covariances of the "posterior" state vector $\alpha_{t-1}|\Im_{t-1}$. After the new data arrives, that is, $\Im_t = Y_t \cup \Im_{t-1}$, we update the prior moments via Equations (5), (7), and (9) by replacing $\hat{\alpha}_{t-1} = a_{t|t-1}$ and $P_{t-1} = B_{t|t-1}$. Then, ignoring the irrelevant constants, we recursively build the log-likelihood function,

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

where $\det(\cdot)$ denotes the determinant. For further details, see Harvey (1994) or Xie et al. (1997).

Next, to estimate the parameter vector $\theta = (\lambda', \gamma', \pi', diag(Q), diag(R))'$ with p elements (as necessary for multiple regions), we maximize Equation (11) so that,

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

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

$$\hat{C} = -\left.\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\hat{\theta}}, \tag{13}$$

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

Moreover, to make inferences robust to mis-specification errors, 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 $L(\theta)$ for each of the T observations. In correctly specified models, $C = V$ and so both the equations (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 $\tilde{C}$. We next conduct simulations to learn how well the proposed approach recovers the model parameters.

4. Monte Carlo Simulations

4.1 Simulation Settings

The simulations shed light on the two issues: the consequences of measurement errors in the *dependent* variables of dynamic models, and the efficacy of the proposed approach. To this end, we set the parameter values, generate multiple data sets, estimate the parameters of the standard and proposed models, and then compare the results with the known parameter values. Specifically, letting T = 100 periods, we generate 1000 data sets using equations (1) and (2) with $\beta_{i,t+1} = \gamma_i z_t$. The true parameters are $\lambda = \gamma_1 = \gamma_2 = 0.5$ and $\pi = 0.1$. The standard deviations of measurement and transition errors equal 100; promotion $v_t \sim U(0,1)$, distribution $z_t \sim 100 \times U(0,1)$, and advertising $u_t \sim 250 \times U(0,1)$, where $U(0,1)$ is the uniform random variable; and the initial $S_0 = 1000$. Applying the proposed approach to the above 1000 simulated data sets, we estimate the model parameters one thousand times. We also estimate the standard marketing-mix model 1000 times using the dependent variable as retail offtakes ($w = 1$), secondary sales ($w = 0$), and composite sales with $w \in [0.1, 0.9]$ in steps of 0.1.

[Insert Table 2 about here]

4.2 Simulation Results

Table 2 presents the estimates from both the proposed and standard models. Recall that the true values are $\lambda = \gamma_1 = \gamma_2 = 0.5$ and $\pi = 0.1$. The proposed approach recovers them satisfactorily: $\hat{\lambda} = 0.484$ (-3.22%), $\hat{\gamma}_1 = 0.510$ (2.00%), $\hat{\gamma}_2 = 0.515$ (3.09%), and $\hat{\pi} = 0.107$ (7.00%). The proposed approach works because it controls for unreliability in the observed metrics.

In contrast, the standard model fails to recover the true parameters. Averaging over all weights, the standard model exhibit severe downward biases: $\hat{\lambda} = 0.254$ (-49.22%), $\hat{\gamma}_1 = 0.273$ (-45.45%), and $\hat{\gamma}_2 = 0.230$ (-54.08%). This downward bias for $\hat{\lambda}$ varies from 17.86% to 98.67%; for $\hat{\gamma}_1$ it varies from 2.20% to 90.05%; for $\hat{\gamma}_2$ it varies from 11.02% to 96.54%. Given that all parameters are under-estimated, the resulting optimal budget will be under-stated. Consequently, managers who rely on the standard model will under-spend on marketing activities.

The severe biases from the standard model reveal that the presence of measurement noise is not innocuous. This finding — noise in the dependent variable induces biases — identifies a boundary condition for errors-in-variable theory, which incorrectly suggests without qualification that noisy dependent variables are innocuous (see, e.g., Greene 1993, Ch. 9). In contrast to the standard theory, noisy *dependent* variables also induce biases *when the models are dynamic*. Next, we apply the proposed approach to real data from Indian markets.

5. Empirical Analysis

5.1 Data Description

We analyze the marketing mix activities and sales outcomes for a major brand of hair care in India. Due to nondisclosure agreements, we cannot divulge the brand's identity and proprietary data. But we note that 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 70 million units are sold monthly nationwide, reaching 130 million consumers (about 25 million households) through a distribution network of over 3 million outlets in India. To provide an empirical generalization, we present results from six political regions that span the breadth of the country: Andhra Pradesh, Gujarat, Karnataka, Maharashtra, Tamil Nadu, and Uttar Pradesh.

The two sales metrics are retail offtakes (i.e., quantity sold by retailers to consumers) and secondary sales (i.e., quantity sold to retailers). The two marketing activities are advertising and promotion over time. 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 exhibits synergies with advertising and promotion. In other words, in the presence of synergy, advertising (or promotion) effectiveness enhances due to wider brand availability: the greater the penetration, the more effective the advertising (or promotion). Table 3 presents the descriptive statistics.

[Insert Table 3 about here]

For each region, we estimate the proposed model in Equations (3) and (4) by combining the metrics as in Proposition 1 and applying the estimation approach described in section 3.3. In addition, we account for potential endogeneity in advertising and promotion using an instrumental variables approach (Bronnenberg and Mahajan 2001, p. 286). We predict each political region's advertising spending using spending in all other regions, and use this predicted spending as the regressor for advertising (e.g., Aravindakshan, Peters, and Naik 2012). Similarly, we predict each political region's promotion timing using the seasonality index from other political regions and other products in the category, and use this predicted promotion timing as the regressor for promotion. We next describe the results.

5.2 Estimation Results

5.2.1 Fit and Forecasts

Table 4 shows the fit and forecast for all six regions. As Table 4 shows, the model fits the data from all six regions satisfactorily. For example, in Maharashtra, the fit for retail offtakes (MAPE = 9.73%) is better than that for secondary sales (MAPE = 17.62%). 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 = 11.62%) is better than that for retail offtakes (MAPE = 14.72%). We next describe the parameter estimates.

[Insert Tables 4 and 5 about here]

5.2.2 Unreliability Estimates

We focus on the estimates of measurement noise and relative bias. Table 5 presents the parameter estimates and robust t-values for all six regions. First, the measurement noise in retail offtakes is large and significant in all the six political regions (Andhra Pradesh: $\sigma_1 = 11.93, t = 6.92$, Gujarat: $\sigma_1 = 8.24, t = 8.29$, Karnataka: $\sigma_1 = 11.08, t = 5.77$, Maharashtra: $\sigma_1 = 18.18, t = 6.34$, Tamil Nadu: $\sigma_1 = 6.35, t = 6.81$, Uttar Pradesh: $\sigma_1 = 7.10, t = 5.52$). Second, the measurement noise in secondary sales also is large and significant in all six political regions (Andhra Pradesh: $\sigma_2 = 27.04,\ t = 14.31$, Gujarat: $\sigma_2 = 21.93, t = 6.30$, Karnataka: $\sigma_2 = 16.20, t = 5.21$, Maharashtra: $\sigma_2 = 69.10, t = 6.99$, Tamil Nadu: $\sigma_2 = 5.05, t = 6.71$, Uttar Pradesh: $\sigma_2 = 5.09, t = 4.51$). Third, the bias in the secondary sales metric is large and significant in all the six political regions (Andhra Pradesh: $\pi_1 = 1.06, t = 20.66$, Gujarat: $\pi_2 = 1.36, t = 13.52$, Karnataka: $\pi_3 = 1.09, t = 19.28$, Maharashtra: $\pi_4 = 1.69, t = 19.14$, Tamil Nadu: $\pi_5 = 0.95, t = 18.27$, Uttar Pradesh:$\pi_6 = 1.09, t = 15.62$). Thus, systematically across the six political regions of India's emerging markets, these results furnish the first empirical evidence that both the metrics are unreliable.

The presence of unreliability renders the parameter estimates inconsistent (Naik and Tsai 2000). In other words, managers will estimate parameters of standard marketing-mix models inaccurately even if the sample size were asymptotically large. In contrast, the proposed approach resolves this problem by filtering out the measurement noise (via Equation 2).

[Insert Table 6 about here]

5.2.3. Carryover Effects and Marketing Elasticities

The consumer demand exhibits strong carryover effects in all the six political regions --- it is large and significant with the median value of 0.9 and ranges from 0.864 to 0.938. Furthermore, there exists synergy between distribution and advertising in some regions. Specifically, advertising effectiveness increases as brand availability increases in Andhra Pradesh ($\gamma_1 = 0.007, t = 4.74$), Karnataka ($\gamma_1 = 0.04, t = 1.97$), and Maharashtra ($\gamma_1 = 0.09, t = 2.64$). We also furnish the first empirical evidence for synergy between distribution and promotion. Specifically, the effectiveness of promotion increases as the distribution intensity increases in Gujarat ($\gamma_2 = 0.76, t = 2.68$), Maharashtra ($\gamma_2 = 0.87, t = 2.02$), Tamil Nadu ($\gamma_2 = 0.49, t = 3.49$), and Uttar Pradesh ($\gamma_2 = 0.41, t = 4.08$).

Next, to interpret the advertising and promotion effects more meaningfully, we compute the elasticity of advertising and promotion. Note that elasticity means one percent change in advertising (or promotion) results in α (or ρ) percentage change in the true consumer demand. Denoting $(\overline{z_\iota}, \overline{u_\iota}, \bar{S_\iota})$ as the mean values of distribution, advertising and consumer demand, respectively, we derive from equation (3) the advertising elasticity $\omega_i = (0.5\gamma_{i1}\overline{z_J}\sqrt{\overline{u_i}})/\, \bar{S}_i$ and the promotion elasticity $\upsilon_i = (\gamma_{i1}\overline{z_\iota})/\, \bar{S}_i$ for the region i. Based on the estimated parameters, we present the elasticities in Table 6. Across the six regions, the mean advertising and promotion elasticity are 0.014 and 0.39, respectively. The estimates of advertising elasticity are smaller than those in the US, where typical it is about 0.10 (Sethuraman, Tellis and Briesch 2011). A direct comparison of promotion elasticity with those in the US is harder to make due to non-availability of meta-analysis. Nonetheless, our mean promotion elasticity is smaller than 2.21 reported in some US settings by Nijs et al. (2001).

Despite these differences in emerging markets, a generalization that holds is that the promotion elasticity exceeds advertising elasticity, indicating that consumers are more responsive to promotional offers than advertising messages. These relative elasticities, together with the costs of marketing activities, guide the allocation of marketing budgets across the regions and activities, which we next discuss.

6. Normative Analysis

Given the unreliability of sales metrics, how should brand managers determine the optimal advertising spending and promotion timing? How should they alter optimal advertising and promotion as unreliability increases? To answer these substantive issues, we formulate and solve a manager's decision-making problem.

6.1. Decision-making Problem

Suppose the manager decides to spend on advertising and promotion over time as follows $\{u_t, v_t : t \in (1,2,\ldots)\}$. Given this marketing-mix plan, the manager generates a sales sequence $S(t)$ measured via two noisy metrics $Y_1(t)$ and $Y_2(t)$, earning an associated profit stream. A manager's decision-making problem is to determine the optimal advertising spending and promotion timing sequence so as to maximize the net present value of profits, given by the objective function $\Pi(u, v)$. The formal problem is as follows:

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

subject to the dynamic sales evolution $\frac{dS}{dt} = \beta_1(t)\sqrt{u(t)} + \beta_2(t)v(t) - \delta S$,

where $\delta = 1 - \lambda$ and $c(t)$ is the cost of promotion.

Since $S(t)$ is not directly observed and rather it is measured via two noisy metrics, we apply Ito's lemma to obtain the stochastic evolution of the observed metrics:

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

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

Thus the presence of 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 complex because, mathematically, the optimal solution to-be-derived has to take into account the following multiple trade-offs: the present versus future (captured through the discount rate ρ), the differential effectiveness of advertising and promotion (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 in each week, and the optimal promotion indicator $v^*(t) \in \{0,1\}$, which informs whether or not to spend on promotion given the time-varying promotional cost $c(t)$. Consequently, even the control domains are mixed: continuous-valued control for advertising and binary switch for promotional timing.

6.2. Optimal Advertising and Promotion with Data Unreliability

Nonetheless, we solve the above stochastic control problem analytically, relegate its proof to Appendix C, and present here the final results. Let us denote $u_0^*(t)$ and $v_0^*(t)$ as the optimal advertising and optimal promotion, respectively, in the presence of perfectly reliable metrics (i.e., with no noise or bias). Then the optimal advertising and promotion in the presence of unreliable metrics are given by

PROPOSITION 2.

$$u^*(t) = u_0^*(t) \times \left(\frac{\phi_1 + \pi^2\phi_2}{1 + \pi(\phi_1 + \phi_2)}\right)^2$$

and

$$v^*(t) = \begin{cases} 1, & v_0^*(t) \times \frac{\phi_1+\pi^2\phi_2}{1+\pi(\phi_1+\phi_2)} > c(t) \\ 0, & \text{otherwise} \end{cases}$$

PROOF. *See Appendix C.*

We designate the expression $\left(\frac{\phi_1+\pi^2\phi_2}{1+\pi(\phi_1+\phi_2)}\right)$ as the correction factor (*CF*). It differs from the optimal weights derived in Proposition 1 and depends on the bias and the signal-to-noise ratios in a non-trivial manner. Moreover, it moderates the optimal advertising and promotion decisions under perfect reliability. Therefore, to quantify it, managers can apply the estimation approach in section 3.3 to their market data and thus incorporate the effects of unreliability in their decision-making.

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

PROPOSITION 3. *Suppose the metrics are unbiased ($\pi = 1$). Then, as unreliability increases, the marketing spending should be reduced. This reduction is more severe for advertising than for promotion.*

PROOF. *CF =* $\frac{\phi_1+\phi_2}{1+\phi_1+\phi_2}$ *<1 when* $\pi = 1$. *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 promotion decreases.*

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 effects of measurement errors. Consequently, they would act as if the metrics are perfect (i.e., noise free), which entails over-spending on advertising and promotion as implied by Proposition 3. In other words, it pays to quantify the magnitude of measurement errors, estimate the signal-to-noise ratios, and then adjust the spending levels as per the correction factor. Overconfidence in data quality is hazardous for 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 promotion, thereby generating the interaction effect.

In sum, this normative analysis demonstrates how unreliable data shapes advertising and promotion decisions.

7. Conclusion

India represents an emerging opportunity for retailers as the middle class expands from 5% to 40% of the population and disposable incomes rise at 5.8% annually. However, an overwhelming portion (over 90%) of the retail market is unorganized with no barcode technology. Consequently marketing managers learn about their brands' sales via unreliable metrics. To mitigate unreliability, managers obtain multiple metrics, which create additional issues of dealing with relative bias in the metrics and optimally combining multiple unreliable metrics. These challenges are pervasive as they exist in 14 million *kirana* shops in India, which means managers cannot directly import models from the West to assess the impact of marketing effectiveness and thus allocate marketing resources optimally.

To address these issues, we present empirical and theoretical analyses of the effects of data unreliability on both the marketing-mix effectiveness and optimal allocations. First, via Proposition 1, we present the optimal weights to combine retail offtakes and secondary sales. This optimal combination not only involves a non-trivial function of relative bias and signal-to-noise ratios of the two sales metrics, but it does not require the normality of measurement errors. Moreover, we show how convex combinations of the metrics yield incorrect elasticities, and how our proposed method correctly recovers advertising and promotion effectiveness. Subsequently, we use real data and validate our model in six different political regions in the country. In the empirical analysis, we find that the measurement noise and relative biases are significant across all the regions. In the normative analysis, we contribute new propositions to the extant literature. Solving the stochastic control problem with mix controls, Proposition 2 derives the optimal advertising and promotion decisions in the presence of unreliability metrics. Our results provide the correction factor that managers quantify using the proposed estimation approach (in section 3.3), and thus adjust their decisions to incorporate the effects of unreliability. Proposition 3 shed light on how overconfidence in data quality result in marketing overspending. In conclusion, we hope that managers use the proposed model and the estimation approach to filter out measurement noises in the metrics and adjust their marketing decisions to incorporate the effects of data unreliability.

Table 1. Elasticity Estimates with Unreliable Data

	Karnataka Region		*Maharashtra Region*	
	Advertising Elasticity	Promotion Elasticity	Advertising Elasticity	Promotion Elasticity
Retail Offtakes (w = 1)	0.017	0.035	0.020	0.012
Secondary Sales (w = 0)	0.005	0.541	0.024	0.618
Composite Sales = w × Retail Offtakes + (1 − w) × Secondary Sales				
w = 0.1	0.003	0.494	0.023	0.58
w = 0.2	0.001	0.446	0.023	0.539
w = 0.3	-0.001	0.397	0.023	0.494
w = 0.4	-0.003	0.348	0.023	0.445
w = 0.5	-0.005	0.297	0.022	0.392
w = 0.6	-0.008	0.247	0.022	0.332
w = 0.7	-0.01	0.195	0.022	0.267
w = 0.8	-0.012	0.142	0.021	0.194
w = 0.9	-0.014	0.089	0.021	0.112

Table 2. Simulation Results

Parameter	λ		γ_1		γ_2	
	Estimate	*% Error*	*Estimate*	*% Error*	*Estimate*	*% Error*
True Values	**0.5**		**0.5**		**0.5**	
	Proposed Model					
	0.484	-3.22	0.510	2.00	0.515	3.09
	Standard Model					
Retail Offtakes (w = 0)	0.007	-98.67	0.050	-90.05	0.017	-96.54
Secondary Sales (w = 1)	0.411	-17.86	0.489	-2.20	0.445	-11.02
w = 0.1	0.060	-87.92	0.094	-81.19	0.059	-88.22
w = 0.2	0.127	-74.66	0.139	-72.28	0.101	-79.85
w = 0.3	0.191	-61.79	0.183	-63.34	0.143	-71.40
w = 0.4	0.247	-50.68	0.228	-54.36	0.186	-62.87
w = 0.5	0.292	-41.63	0.273	-45.36	0.229	-54.28
w = 0.6	0.328	-34.44	0.318	-36.34	0.272	-45.65
w = 0.7	0.356	-28.77	0.363	-27.31	0.315	-37.00
w = 0.8	0.379	-24.28	0.409	-18.28	0.358	-28.34
w = 0.9	0.396	-20.72	0.454	-9.24	0.402	-19.68
Average	**0.254**	**-49.22**	**0.273**	**-45.45**	**0.230**	**-54.08**

Table 3. 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

Table 4. Fit and Forecast

	Andhra Pradesh	Gujarat	Karnataka	Maharashtra	Tamil Nadu	Uttar Pradesh
Retail Offtake Fit (MAPE)	8.03	14.89	13.45	9.73	23.33	33.20
Retail Offtake Forecast (MAPE)	5.90	18.77	13.17	9.47	14.72	18.68
Secondary Sales Fit (MAPE)	14.84	23.27	15.61	17.62	16.82	19.45
Secondary Sales Forecast (MAPE)	19.79	24.3	12.7	23.41	11.62	17.30

Table 5. Estimation Results

	Andhra Pradesh		**Gujarat**		**Karnataka**		**Maharashtra**		**Tamil Nadu**		**Uttar Pradesh**	
Parameters	*Estimate*	*t-value*	*Estimate*	*t-value*	*Estimate*	*t-value*	*Estimate*	*t-value*	*Estimate*	*t-value*	*Estimate*	*t-value*
Carryover Effect, λ	**0.938**	**60.189**	**0.924**	**26.721**	**0.909**	**22.718**	**0.879**	**26.498**	**0.895**	**27.742**	**0.864**	**23.830**
Distribution Synergy with Advertising, γ_1	**0.007**	**4.743**	0.001	0.014	**0.004**	**1.968**	**0.009**	**2.642**	0.001	0.294	0.001	0.570
Distribution Synergy with Promotion, γ_2	**0.251**	**1.821**	**0.757**	**2.682**	0.374	1.329	**0.876**	**2.020**	**0.497**	**3.489**	**0.412**	**4.087**
Relative Bias, π	**1.063**	**20.658**	**1.360**	**13.517**	**1.093**	**19.277**	**1.695**	**19.138**	**0.944**	**18.267**	**1.095**	**15.619**
Retail Offtakes Noise, σ_1	**11.928**	**6.921**	**8.237**	**8.282**	**11.085**	**5.763**	**18.180**	**6.343**	**6.354**	**6.810**	**7.102**	**5.529**
Secondary Sales Noise, σ_2	**27.044**	**14.305**	**21.932**	**6.287**	**16.155**	**5.208**	**69.086**	**6.985**	**5.053**	**6.706**	**5.098**	**4.509**

***Bold** estimates are statistically significant at the 95% confidence level.

Table 6. Elasticity Estimates in India's Emerging Markets

	Andhra Pradesh	Gujarat	Karnataka	Maharashtra	Tamil Nadu	Uttar Pradesh
Advertising Elasticity	0.029	0.001	0.020	0.032	0.001	0.001
Promotion Elasticity	0.119	0.409	0.249	0.335	0.513	0.712

APPENDIX A: OBTAINING MATRIX P_t TO COMBINE METRICS OPTIMALLY

We derive Equation (7) by noting that $P_t = E[\nu_t \nu_t']$, where

$$\begin{aligned}
\nu_t &= \alpha_t - \hat{\alpha}_t \\
&= \alpha_t - (\hat{\alpha}_{t-1} + K_t(Y_t - \hat{Y}_t)) \\
&= (\alpha_t - \hat{\alpha}_{t-1}) - K_t(Y_t - H\hat{\alpha}_{t-1}) \\
&= \nu_{t-1} - K_t(H\alpha_t + \varepsilon_t - H\hat{\alpha}_{t-1}) \\
&= \nu_{t-1} - K_t H(\alpha_t - \hat{\alpha}_{t-1}) - K_t\varepsilon_t \\
&= \nu_{t-1} - K_t H\nu_{t-1} - K_t\varepsilon_t \\
&= (I - K_t H)\nu_{t-1} - K_t\varepsilon_t .
\end{aligned}$$

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

$$\begin{aligned}
P_t &= E[\nu_t \nu_t'] \\
&= E[\{(I - K_t H)\nu_{t-1} - K_t\varepsilon_t\}\{(I - K_t H)\nu_{t-1} - K_t\varepsilon_t\}'] \\
&= (I - K_t H)E[\nu_{t-1}\nu_{t-1}'](I - K_t H)' - (I - K_t H)E[\nu_{t-1}\varepsilon_t']K_t' - K_t E[\varepsilon_t \nu_{t-1}'](I - K_t H)' + K_t E[\varepsilon_t \varepsilon_t']K_t' \\
&= (I - K_t H)P_{t-1}(I - K_t H)' + K_t R K_t',
\end{aligned}$$

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 optimally combine the unreliable metrics Y_1 and Y_2.

Recall that $\hat{S} = k_1^* Y_1 + k_2^* Y_2$, where $[k_1^* \quad k_2^*]$ are elements in the first row of the matrix K_t^*. We know that P_{t-1} is the variance of the true consumer demand σ_S^2, H represents the bias vector $[1 \quad \pi]'$, and R represents the diagonal variance matrix $\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$. Using these values, we compute K_t^* as follows:

$$
\begin{aligned}
K_t^* &= P_{t-1}H'(HP_{t-1}H' + R)^{-1} \\
&= \sigma_S^2 [1 \quad \pi](HP_{t-1}H' + R)^{-1} \\
&= \sigma_S^2 [1 \quad \pi] \begin{bmatrix} \pi\sigma_S^2 + \sigma_1^2 & \sigma_S^2 \\ \pi\sigma_S^2 & \pi\sigma_S^2 + \sigma_2^2 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} \frac{\sigma_S^2\sigma_2^2 + \pi(1-\pi^2)\sigma_S^4}{\sigma_1^2\sigma_2^2 + \pi\sigma_S^2(\sigma_1^2+\sigma_2^2)} & \frac{\pi\sigma_S^2\sigma_1^2 + \pi(\pi^2-1)\sigma_S^4}{\sigma_1^2\sigma_2^2 + \pi\sigma_S^2(\sigma_1^2+\sigma_2^2)} \end{bmatrix}.
\end{aligned} \tag{B1}
$$

In (B1), $k_1^* = \frac{\sigma_S^2\sigma_2^2 + \pi(1-\pi^2)\sigma_S^4}{\sigma_1^2\sigma_2^2 + \pi\sigma_S^2(\sigma_1^2+\sigma_2^2)}$ and $k_2^* = \frac{\pi\sigma_S^2\sigma_1^2 + \pi(\pi^2-1)\sigma_S^4}{\sigma_1^2\sigma_2^2 + \pi\sigma_S^2(\sigma_1^2+\sigma_2^2)}$. To further simplify these expressions, we define $\phi_1 = \frac{\sigma_S^2}{\sigma_1^2}$ *and* $\phi_2 = \frac{\sigma_S^2}{\sigma_2^2}$ to be the signal-to-noise ratios of Y_1 and Y_2, respectively. Dividing the numerator and denominator of k_1^* by $\sigma_1^2\sigma_2^2$, we express k_1^* as a function of the bias and signal-to-noise ratios as

$$k_1^* = \frac{\phi_1 + \pi(1-\pi^2)\phi_1\phi_2}{1+\pi(\phi_1+\phi_2)}. \tag{B2}$$

Similarly, dividing the numerator and denominator of k_2^* by $\sigma_1^2\sigma_2^2$, we obtain

$$k_2^* = \frac{\pi\phi_2 + \pi(\pi^2-1)\phi_1\phi_2}{1+\pi(\phi_1+\phi_2)}. \tag{B3}$$

The closed-form expressions in (B2) and (B3) comprise the optimal weights by which to combine the multiple unreliable metrics, thus proving Proposition 1.

APPENDIX C: DERIVATION OF THE OPTIMAL ALLOCATIONS WITH UNRELIABLE DATA

We seek to solve the marketing mix allocation problem stated in Equation (15):

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

subject to

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

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

Applying Ito's lemma, we observe that $dY_1 = dS + \sigma_1 dW_1$, where $W_1(t)$ is the standard Wiener process. Then, $dS = [\beta_1(t)\sqrt{u(t)} + \beta_2(t)v(t) - \delta(Y_1(t) - \varepsilon_1(t))]dt = [\beta_1(t)\sqrt{u(t)} + \beta_2(t)v(t) - \delta Y_1]dt + \delta\sigma_1 dW_1$. Next, by substituting this expression for dS in $dY_1 = dS + \sigma_1 dW_1$, we get equation (16), which represents the sales dynamics in the observed metric Y_1 . Similarly, we derive the equation (17).

We note that the presence of measurement noise in the metrics introduces uncertainty, which is represented by the Wiener processes. Consequently, to maximize the total profit $\Pi(u, v)$, we need to solve a stochastic control problem. To this end, we formulate the stochastic Hamilton-Jacobi-Bellman equation as follows:

$$\rho V = Max[(m(k_1^* Y_1 + k_2^* Y_2) - u - cv) + \dot{V}_1\big(\beta_1\sqrt{u} + \beta_2 v - \delta Y_1\big) + \dot{V}_2\big(\pi\beta_1\sqrt{u} + \pi\beta_2 v - \delta Y_2\big) + 0.5\sigma_1^2\ddot{V}_1 + 0.5\sigma_2^2\ddot{V}_2], \quad \text{(C1)}$$

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 Equation (C1) is a second-order partial differential equation.

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

$$-1 + \frac{\dot{V}_1\beta_1}{2\sqrt{u}} + \frac{\dot{V}_2\pi\beta_1}{2\sqrt{u}} = 0, \quad \text{(C2)}$$

which upon re-arrangement gives the optimal advertising:

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

Based on previous research (e.g., Aravindakshan, Peters, Naik 2012), we conjecture and confirm that the value function $V(Y_1, Y_2) = v_0 + v_1 Y_1 + v_2 Y_2$ satisfies the partial differential equation (C1). 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 $u^* = (0.5\beta_1(v_1 + \pi v_2))^2$ in the stochastic HJB equation (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:

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

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 (C1) with respect to *v* to get

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

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 (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:

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

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

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

References

Aravindakshan, Ashwin, Kay Peters, and Prasad A. Naik (2012), "Spatiotemporal Allocation of Advertising Budgets," *Journal of Marketing Research,* 49 (1), 1-14.

Bronnenberg, Bart J. and Vijay Mahajan (2001), "Unobserved Retailer Behavior in Multimarket Data: Joint Spatial Dependence in Market Shares and Promotion Variables," *Marketing Science,* 20 (3), 284–99.

Greene, William H. (1993), *Econometric Analysis*. New York, NY: Macmillan Publishing Company.

Hanssens, Dominique, Leonard Parsons, and Randall Schultz (2001), *Market Response Models: Econometric and Time Series Analysis*, 2nd ed. Boston, Mass.: Kluwer Academic Publishers.

Harvey, Andrew C. (1994), *Forecasting, Structural Time Series Models and the Kalman Filter*. New York, N.Y.: Cambridge University Press.

Kolsarici, Ceren, and Demetrios Vakratsas (2010), "Category Versus Brand-Level Advertising Messages in a Highly Regulated Environment," *Journal of Marketing Research,* 47(6), 1078-1089.

Leone, Robert P. (1995), "Generalizing What Is Known About Temporal Aggregation and Advertising Carryover," *Marketing Science,* 14(3), G141-G150.

McKinsey Global Institute Report, "The Bird of Gold: The Rise of India's Consumer Market," May 2007, McKinsey & Company.

Naik, Prasad A., and Chih-Ling Tsai (2000), "Controlling Measurement Errors in Models of Advertising Competition," *Journal of Marketing Research,* 37(1), 113-24.

Naik, Prasad A., and Kalyan Raman (2003), "Understanding the Impact of Synergy in Multimedia Communications," *Journal of Marketing Research*, 40(4), 375–88.

Narayanan, Sridhar, Ramarao Desiraju, and Pradeep K Chintagunta (2004), "Return on Investment Implications for Pharmaceutical Promotional Expenditures: The Role of Marketing-Mix Interactions," *Journal of Marketing*, 68(3), 90-105.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E.M. Steenkamp, and Dominique M. Hanssens (2001), "The Category-Demand Effects of Price Promotions," *Marketing Science*, 20(1), 1-22.

Schultz, Don E. (2012), "Bad Data, Bad Models or Bad Managers," *Marketing News,* April 30th, 2012, p. 13.

Schumpeter, David (2011), "Indian Retail Wholesale Reform," *Economist,* Nov 25th issue.

http://www.economist.com/blogs/schumpeter/2011/11/indian-retail. Accessed June 28th, 2012.

Sethuraman, Raj, Gerard J. Tellis, and Richard A. Briesch (2011), "How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities," *Journal of Marketing Research*, 48(3), 457-471.

Simon, Julian L., and Johan Arndt (1980), "The Shape of the Advertising Function," *Journal of Advertising Research*, 20(3), 11-28.

White, Halbert (1982), "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1-25.

Xie, Jinhong, Michael Song, Marvin Sirbu, and Qiong Wang (1997), "Kalman Filter Estimation of New Product Diffusion Models," *Journal of Marketing Research,* 34 (3), 378-93.