Evaluating the Effectiveness of Marketing Expenditures

by

Thomas Otter
Fisher College of Business
Ohio State University
otter_2@cob.osu.edu

Ling-Jing Kao
Leavey School of Business
Santa Clara University
LKao@scu.edu

Chih-Chou Chiu
Institute of Commerce Automation and Management
National Taipei University of Technology
chih3c@ntut.edu.tw

Timothy J. Gilbride
Mendoza College of Business
University of Notre Dame
tgilbridi@nd.edu

Greg M. Allenby
Fisher College of Business
Ohio State University
allenby_1@cob.osu.edu

February, 2007
Evaluating the Effectiveness of Marketing Expenditures

Abstract

Marketing expenditures in the form of pricing, product development, promotion and channel development are made to maximize return on investment. A challenge in evaluating the effectiveness of these expenditures is dealing with the potential simultaneity of an output measure, such as sales, and input measures, such as promotional spending, that are jointly determined by consumer preferences and sensitivities. While marketing control variables are explanatory of sales, they are not determined independent of the marketplace. This paper proposes an approach to dealing with the simultaneous relationship among input and output variables that yields measures of the efficiency of converting inputs to outputs, and the optimal allocation of input variables. We illustrate our approach using data from a services company operating in multiple geographic regions.

Keywords: Bayesian Analysis, Random Effects, Technical Efficiency, Allocative Efficiency
1. Introduction

Marketing expenditures in support of product, promotion, and channel activities are made to maximize return on investment by responding to individual wants, attracting new customers and increasing sales. Expenditures are effective when they are allocated across decision options in a manner that leads to maximum value of these criteria, subject to various constraints placed by the firm. Decision options might take the form of different product development efforts, geographic allocation of promotional dollars, and channel-related decisions such as the deployment of a sales force. While, intuitively, it makes sense to allocate greater expenditure to options that are more responsive, the true responsiveness of the options is often difficult to assess, particularly when there is limited disaggregate-level data. Statistical inference must therefore deal with the simultaneous nature of expenditures on marketing variables (i.e., the 4P’s) and responses that are jointly driven by common response coefficients.

Consider, for example, units described by a simple regression model \( y_i = x_i' \beta + \epsilon_i \) where "i" is the unit index, y is the dependent variable, x denotes expenditure, \( \beta \) is the response coefficient and \( \epsilon \) is error. Assume further that management makes greater allocation to units that are more responsive (i.e., \( x_i \) is large when \( \beta_i \) is large), and that there is insufficient data to estimate the model at the unit-level. We may have just one, or a limited number of observations for any one unit. Simple pooling of data across units results in an aggregate-level regression model \( y = x' \bar{\beta} + u \) with a common regression coefficient \( \bar{\beta} \), and errors \( u = x' (\beta - \bar{\beta}) + \epsilon \) that are positively correlated with the input variable because of the allocation rule. Positive correlation between the error and the explanatory variable invalidates the use of standard estimators – i.e., \( \text{plim} (\hat{\beta}) \neq \text{plim} (\bar{\beta}) \), requiring...
the development of more formal models for inference. We show below that pooling with a random-effects model also has the same problem – i.e., the allocation rule needs to be incorporated into the model.

Endogenously determined variables are commonly encountered in marketing. They occur whenever expenditures are set with the objective of maximizing return on investment, and the level of expenditure is related to the level of response. Conducting analysis for such data requires two things. The first is a measure of the responsiveness of the decision option to levels of expenditure. Response measures are obtained by modeling the association of a dependent (output) variable, y, such as sales, to the variables being allocated (x), also known as the input variables. The second is a model of expenditure allocation. This second item involves an analysis where the inputs become the dependent variable, and the response coefficients are explanatory.

In this paper, we propose a direct method to modeling the simultaneous relationship between demand-side responses to marketing expenditures, and their supply-side allocation. Our approach is potentially applicable whenever explanatory variables are endogenously determined in anticipation of expected returns. By direct, we mean that our approach is not based on the dual cost formulation pursued in the economics literature (e.g., Kumbhakar and Tsionas, 2005), and is therefore applicable to production functions that do not have a tractable dual representation. Our method facilitates inference about the efficiency with which firms allocate and deploy their resources. We demonstrate its use with a variety of response functions encountered in marketing applications.

The organization of the paper is as follows. We introduce our model in the next section and discuss its application to evaluating marketing expenditures. Measures of efficiency are derived and compared to existing measures. In section 3, a simulation study is presented that investigates the effects of endogeneity. Section 4 describes data from a service firm used to illustrate our method.
The firm allocates promotional expenditures to 21 geographic areas to generate new business in these markets, with each allocation spent, or deployed, over a 14 month period. Alternative response models and associated supply-side allocation equations are derived and fit to the data. Implications for assessing efficiencies are then discussed in section 5, and concluding remarks are offered in section 6.

2. Assessing Technical and Allocative Efficiencies

Marketing expenditures are an investment that usually results in a positive return. The return can be in the form of increased sales, or customers, or some form of infra-structure that makes acquiring these items easier in the future – i.e., some expenditures produce an immediate return, while others make it easier to produce future returns. Price promotions and advertising, for example, may have an immediate impact on sales, while some forms of channel investment (e.g., new distribution outlets) may work more on an interactive level to support promotional spending. It is useful to think of marketing investment expenditures as inputs to a hierarchical production function:

\[ y = f(x, \beta) = f(x_1, \beta_1 = g(x_2, \beta_2 = h(x_3, \beta_3 = ...)) \]  

where \( x_1 \) are expenditures whose variations lead to direct effects on \( y \), \( x_2 \) are expenditures associated with \( \beta_1 \) (the response coefficient for \( x_1 \)), and \( x_3 \) are expenditures associated with \( \beta_2 \) (the response coefficient for \( x_2 \)). For example, consider the case where \( x_1 \) denotes expenditures related to the formulation of the good in terms of attributes and benefits, \( x_2 \) denotes advertising expenditures resulting in greater exposure (GRP), and \( x_3 \) reflects the intensity of distribution (ACV). Equation (1) implies that distribution expenditures (\( x_3 \)) serve to make an offering more readily available to a market target, so that advertising expenditures (\( x_2 \)) will have a greater impact. In turn, by alerting
prospects to the attributes and benefits of an offering, the effects of product enhancements will be larger.

The form of the response functions, f, g and h, is general but needs to have specific properties for globally optimal and interior solutions to exist (see Varian, 1992). These conditions are used to create a norm, or standard, by which current expenditures are evaluated. In particular, they must be i) positively valued, ii) have positive marginal returns that iii) diminish as the allocation increases (i.e., diminishing marginal returns). To illustrate, consider a monopolist pricing problem using a constant elasticity model, where it is assumed that the variation in prices over time is due to stochastic departures from optimal price-setting behavior. The likelihood for the data is a combination of a traditional demand model:

$$\ln y_t = \beta_0 + \beta_1 \ln p_t + \epsilon_t; \quad \epsilon_t \sim Normal(0, \sigma^2_{\epsilon})$$

and a factor for the endogenous price variable. Optimal pricing for the monopolist can be shown to be:

$$p_t = mc \left( \frac{\beta_1}{1 + \beta_1} \right) e^{\nu_t}; \quad \nu_t \sim Normal(0, \sigma^2_\nu)$$

where "mc" denotes the marginal cost of the brand, and a supply-side error term has been added to account for temporal variation of observed prices from the optimal price. Taking logs of equation (3) yields:

$$\ln p_t = \ln mc + \ln \left( \frac{\beta_1}{1 + \beta_1} \right) + \nu_t; \quad \nu_t \sim Normal(0, \sigma^2_\nu)$$

The supply-side equation (4) can be used to help locate the value of $\beta_1$ if marginal cost is known. It also imposes ordinal constraints on the model parameters – i.e., $\beta_1 < -1$ for the supply equation to be valid. If $-1 \leq \beta_1 < 0$, then the quantity $\beta_1/(1+\beta_1)$ is negative and the logarithm in equation (4) is not defined. Optimal pricing behavior for the monopolist only exists when self-price effects are
The supply side equation influences estimation by assuming that optimal pricing behavior exists, even when marginal costs are not known. Moreover, the supply-side error variance ($\sigma^2_\nu$) provides a means of assessing allocative (i.e., pricing) inefficiency.

A limitation of the dual cost approaches prevalent in the economic literature is that they require knowledge of costs for analysis to proceed. In contrast, our method results in improved efficiency of estimates even when costs are not known. If input sensitivities ($\beta$) vary across units of analysis (e.g., time, regions, consumers), then, as illustrated earlier, the absence of supply-side information results in inconsistent parameter estimates.

There are many challenges to employing current economic models for assessing the effectiveness of marketing expenditures. These include the facts that i) input prices are often not observed; ii) allocation budgets include capital and expense items; iii) outputs in marketing can be in the form of sales, customers, leads, memory structure, intentions, and consideration sets; iv) expenditures affect both the stock and flow of output (e.g., current and new customers); and v) budget allocation and deployment are not simultaneous (e.g., yearly budgets and monthly expenditures). Each of these items is difficult to accommodate in models that rely on a dual cost specification – i.e., they do not lead to an explicit dual representation. We show below that our method can deal with the above issues while incorporating heterogeneous response coefficients for assessing technical efficiency, and a supply-side model specification for assessing allocative efficiency.

We introduce our method using a simple Cobb-Douglas production function, and then later expand analysis with other production functions for marketing. Many researchers in marketing have used variants of the Cobb-Douglas production function to assess technical and allocative efficiency. Technical efficiency measures the relative ability of a production unit to convert a given set of inputs
to outputs (see Koop, Osiewalski and Steel 1994), and allocative efficiency measures the optimality of the input levels.

Examples of the use of the Cobb-Douglas function in marketing include Misra (2005) who posits a hierarchical model structure to account for district- and regional-effects in decomposing salesforce technical effects, Horsky and Nelson (1995) who proposes an method for benchmarking sales response, Carroll, Green and DeSarbo (1979) who investigate preferences for leisure time, Morey and McCann (1983) who study optimal lead generation from advertising, and Mantrala, Sinha and Zoltners (1992) who compare allocation rules for various concave functions to top-down budgeting practices at firms. None of this literature, however, acknowledges the possibility that if a firm allocates inputs with even an implicit understanding of market response, then both input and output variables are dependent and determined from within the system of study. Thus, these previous methods that condition on the inputs are prone to misspecification due to the presence of simultaneity.

Technical Efficiency

Models of technical efficiency augment equation (1) with a unit-specific parameter, \( \nu_i \), that measures the efficiency of the unit: \( y_i = f(x_i, \beta, \nu_i) \), where \( \nu_i \) is constrained to be greater than zero. In the economics literature, \( \nu_i \) is usually specified as \( y_i = f \left( x_i e^{-\nu_i}, \beta \right) \) and \( \nu_i \) is interpreted as the amount that a technically inefficient producer overuses the inputs compared with an efficient producer producing the same output (Kumbhakar and Tsionas 2005). In marketing, the inefficiency parameter is sometimes introduced as a scalar \( y_i = f \left( x_i, \beta \right) e^{-\nu_i} \) and interpreted as the reduction in output for the same level of input (Dutta, Kamakura and Ratchford 2004). The difference between these measures amounts to whether efficiency is measured in terms of inputs for a given level of
output (i.e., \( y_i = f(x_i e^{-\nu_i}, \beta) \)), where \( \nu_i \) is associated with each element of the input vector \( x_i \), or associated with the output for given levels of input (i.e., \( y_i = f(x_i, \beta e^{-\nu_i}) \)).

Rather than pursue estimates of technical efficiency through the scalar value \( \nu_i \), we allow for output variation for a given level of input by specifying heterogeneous production functions. The advantage of this approach is that some units will be more efficient translating some marketing inputs into outputs than others. We therefore allow the response coefficients to be unit specific, \( \beta_i \). The ability to translate specific marketing inputs into marketing outputs may be outside the control of local management. For instance, promotional spending may be more effective for retailers in geographic regions with high population density than for those with low population density. To the extent that marketing production functions can be heavily influenced by exogenous factors, we allow for heterogeneity in production functions and focus on allocative efficiency.

### Allocative Efficiency

The optimal allocation of inputs occurs when the marginal effect of an additional input unit per unit cost (e.g., dollar of expenditure) is equal across the decision units. Optimal levels \( \langle x^* \rangle \) can be determined by forming the auxiliary production function \( (L) \) that includes the budget constraint and Lagrangian multiplier \( (\lambda) \), taking derivatives, and solving the system of equations arising from the identities \( \partial L / \partial x_i = \partial L / \partial x_k = \partial L / \partial \lambda = 0 \) for \( x^* \). These equations lead to the allocation rule that marginal output divided by input prices \( (p) \) should be the same across inputs:

\[
\frac{\partial y / \partial x_j}{p_j} = \frac{\partial y / \partial x_k}{p_k} \quad \text{for all } j \text{ and } k
\]

(5)

where the subscripts "j" and "k" refer to different input variables, and \( p_j \) denotes the price of input \( j \). Equations (1) and (5) can be used to specify a system of equations where outputs \( (y) \) and inputs \( (x) \)
are jointly determined. Observed deviations from the optimal allocations are accommodated by introducing errors \((\zeta)\) into the price ratio, i.e., \(p_j^* = p_j e^{\zeta_j}\), and assuming that equation (5) holds exactly for \(p_j\) replaced by \(p_j^*\). Estimates of error realizations \(\zeta_i\) can then be used to infer allocative inefficiencies.

Schmidt and Lovell (1979) first used the system consisting of the production function and the first order conditions of cost minimization in the context of a Cobb-Douglas production function. However, the generalization to more flexible productions functions like the translog have proven to be difficult (see Kumbhakar 1989). An alternative method employs the dual cost function formulation. Maximizing output subject to a budget constraint can be shown to be dual to the minimizing costs subject to an output constraint (see Deaton and Muehlbauer 1980). When costs and input prices are observed, estimation is simplified for some production functions such as the translog function (see Kumbhakar and Tsionas, 2005).

Our method deviates from this literature by directly estimating the system of equations formed by the production function and associated first order conditions. To illustrate, consider a standard Cobb-Douglas production function and associated first-order conditions:

\[
y = \beta_0 x_1^{\beta_1} x_2^{\beta_2}
\]

\[
\lambda = \frac{\partial y}{\partial x_1} \frac{p_1}{p} = \frac{\beta_1}{p_1} \beta_0 x_1^{\beta_1 - 1} x_2^{\beta_2}
\]

\[
\lambda = \frac{\partial y}{\partial x_2} \frac{p_2}{p} = \frac{\beta_2}{p_2} \beta_0 x_1^{\beta_1} x_2^{\beta_2 - 1}
\]

Taking logs and rearranging terms we have:

\[
\ln(y) = \ln \beta_0 + \beta_1 \ln(x_1) + \beta_2 \ln(x_2)
\]

\[
\ln(x) = \begin{bmatrix} \ln(x_1) \\ \ln(x_2) \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 \\ \beta_1 & \beta_2 - 1 \end{bmatrix}^{-1} \begin{bmatrix} \ln \lambda - \ln \beta_0 - \ln \beta_1 + \ln p_1 \\ \ln \lambda - \ln \beta_0 - \ln \beta_2 + \ln p_2 \end{bmatrix}
\]
Adding heterogeneity and observation error ($\varepsilon_i$) to the output equation, and assuming inputs are allocated based on $p_j^* = p_j e^{\varepsilon_j}$ rather than $p_j$ leads to estimation equations:

$$\ln(y_i) = \ln(\beta_{yi}) + \beta_{yi} \ln(x_{yi}) + \beta_{zi} \ln(x_{zi}) + \varepsilon_i$$

$$\ln(x_i) = \begin{bmatrix} \ln(x_{yi}) \\ \ln(x_{zi}) \end{bmatrix} = \begin{bmatrix} \beta_{yi} & \beta_{zi} \\ \beta_{yi} & \beta_{zi} \end{bmatrix}^{-1} \begin{bmatrix} \ln \lambda - \ln \beta_{yi} - \ln \beta_{zi} + \ln p_{yi} + \zeta_{yi} \\ \ln \lambda - \ln \beta_{yi} - \ln \beta_{zi} + \ln p_{zi} + \zeta_{zi} \end{bmatrix} \quad (8)$$

This system is just identified for known $\lambda$ and prices, and repeated observations over time allow estimation of all model parameters. When input prices are missing, the assumption of constant input prices identifies the allocation errors, alternatively, prior restrictions on the allocation errors can identify unobserved input prices. In this system, variation in the input variables is entirely due to the allocation errors, similar to the monopolist pricing example discussed earlier.

The allocation equations constrain the parameter space to the subspace for which global interior optima, i.e. where it is optimal to allocate positive inputs to more than one unit, are obtained. Across units $i$ the observed input allocations are a function of the unobserved response coefficients which is critical for any type of analysis that attempts to pool information across units $i$. In section 3 we develop a model where allocation decisions are made on a time scale different from the decision when to deploy how much of the pre-allocated budget. Recently, Kumbhakar and Wang (2006) propose a similar approach but specify the first-order conditions in terms of cost shares, not inputs ($x$). Their method also does not allow for heterogeneity in the production units.

Our method is related to the empirical industrial organization literature that examines the simultaneity of prices in models of supply and demand (Hausman 1996, Villas-Boas and Winer 1999, Nevo 2001, Yang, Chen and Allenby 2003). In these models, the supply-side equation has prices as dependent variables in models that assume optimal behavior of the firm, where deviation from optimality is due entirely to supply-side shocks that affect marginal costs. These assumptions are
made so that alternative competitive assumptions (e.g., Nash equilibrium) can be tested. In contrast, our production function specification allows for departures from optimality that facilitate assessment of the degree that expenditures are sub-optimally allocated, and allows for multiple input variables allocated across multiple geographic regions.

Our method also builds on recent approaches to modeling strategically determined input variables (Manchanda, Rossi and Chintagunta 2004) that employ descriptive functions, as opposed to production functions such as equation (1). Descriptive response functions are not necessarily associated with a priori sensible, or globally optimal, allocation rules, such as those derived in equations (6) – (8) for the Cobb-Douglas model, and therefore cannot be used to provide allocation norms. However, descriptive functions can be used to assess the effects of incremental changes to the input variables, i.e., whether an increases in an input lead to an increased output. Our method allows for the simultaneous assessment of allocation across all units and ensures the presence of an optimal solution by using response functions that have a global maximum. Thus, it is better suited for making global assessments of input allocation.

**Bayesian Estimation**

Bayesian estimation of the system of equations in (8) proceeds by recursively generating draws from the full conditional distribution of all model parameters (see Rossi, Allenby and McCulloch 2005, chapter 7). The likelihood for the data can be written as:

$$\pi \left( \ln y, \{ \ln x_k \} | \beta, \lambda, \sigma^2, \Sigma \right) = \pi \left( \ln y | \{ \ln x_k \}, \beta, \sigma^2 \right) \times \pi \left( \{ \ln x_k \} | \beta, \lambda, \Sigma \right)$$  \hspace{1cm} (9)

The first factor on the right corresponds to the production function likelihood where the input variables (x) are treated as independent variables. The second factor corresponds to the system of allocation equations. The likelihood for these equations require a Jacobian term because of
simultaneity. For the system of allocation equations corresponding to the Cobb-Douglas system in equation (8), we have:

\[
\pi \left( \{ \ln x_k \} \mid \beta, \lambda, \Sigma \zeta \right) = \pi \left( \zeta \mid \Sigma \right) \cdot J_{\zeta \rightarrow \ln x}
\]

(10)

where

\[
\zeta = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} = \begin{bmatrix} \beta_1 - 1 & \beta_2 \\ \beta_2 - 1 & \beta_1 - 1 \end{bmatrix} \begin{bmatrix} \ln (x_1) \\ \ln (x_2) \end{bmatrix} + \begin{bmatrix} -\ln \lambda + \ln \beta_0 + \ln \beta_1 - \ln p_1 \\ -\ln \lambda + \ln \beta_0 + \ln \beta_2 - \ln p_2 \end{bmatrix}
\]

(11)

and

\[
J_{\zeta \rightarrow \ln x} = (\beta_1 - 1)(\beta_2 - 1) - \beta_1 \beta_2
\]

(12)

\(\Sigma\zeta\) are parameters associated with the distribution of the allocative errors (e.g., \(\zeta \sim N(0, \Sigma\zeta)\)).

Equation (11) places ordinal constraints on the likelihood for the coefficients \(\beta_1\) and \(\beta_2\) in two ways. First, the coefficients must be positive for the logarithm to exist. Second, the Jacobian of the likelihood is zero when \(\beta_1 + \beta_2 = 1.0\). These constraints, coupled with the observed inputs \((x)\) needing to be positive for the function \(\ln(x)\), results in a constrained parameters space \(0 < \beta_1 + \beta_2 < 1.0\) that imposes useful information for estimation, even when input prices are not observed and the prior on the allocative error, \(\pi(\zeta)\), is flat. An advantage of conducting Bayesian estimation is that, given the model parameters, the evaluation of these components of the likelihood is straightforward and estimation can proceed using the Metropolis algorithm without the need to compute gradients and the Hessian of the likelihood. The supply-side equations in (11) facilitate the analysis of allocative efficiency, and, as shown below, heterogeneity is introduced to allow for differences in relative efficiency.
3. Simulation Study

The efficiency literature in economics typically assumes a production process where current output is distinct from past outputs, or units in stock. In marketing, the process by which output is generated does not generally allow for this distinction. Many marketing actions have an effect on stock (e.g., the likelihood of retaining a current customer) and on production (e.g., the number of new customers attracted). Thus, output for many problems in marketing needs to be defined more broadly, possibly as the net increase in the number of new "products" (e.g., customers) per time period.

Another distinguishing feature of measuring outputs in marketing is that an input variable can take on a value of zero (e.g., no promotional expenditure in a given month). However, zero input expenditures do not necessarily translate into zero outputs. Output may be non-zero for a number of reasons, such as the lagged effect of the inputs and the possibility of omitted variables in the analysis. Output may also be negative because of competitive influences, poor marketplace execution, and competitive practices. Thus, marketing production functions often need to be specified with additive intercepts, lagged input variables and an additive error term.

Finally, market budget allocation decisions are often made on a yearly basis, while expenditures and their effects may take place monthly. Allocation decisions for some inputs, such as physical plant, may be subject to a different budget constraint than operational expenditures, yet both may be important in generating output. The specification of marketing production functions may need to allow for different levels of temporal aggregation, and possibly different budget constraints ($\lambda$) for the different input variables.

Effects of endogeneity are investigated using a model similar to that employed in our empirical analysis below. The model is sufficiently general to apply to a wide variety of settings in marketing. We demonstrate that inferences based only on demand-side equations yield estimates of
response coefficients that are inconsistent, similar to the simple regression model described in the introduction. We begin by assuming there are N geographic areas indexed by i, i = 1, …, N that produce an output at time t, t = 1, …, T according to:

$$y_{i,t} = \beta_{0,i} + \gamma_i x_{1,i,t}^\beta_{1,i} x_{2,i,t}^\beta_{2,i} + \epsilon_{i,t}$$ (13)

The inputs are allocated across areas based on an implicit understanding of the productivity of individual areas such that the joint output of the system is maximized. At each time point, in each area, the first order conditions are

$$\lambda_i = \frac{dy_{i,t}}{p_{1,i}dx_{1,i,t}} = \frac{1}{p_{1,i}} \gamma_i \beta_{1,i} x_{1,i,t}^{\beta_{1,i}-1} x_{2,i,t}^\beta_{2,i}$$

$$\lambda_2 = \frac{dy_{i,t}}{p_{2,i}dx_{2,i,t}} = \frac{1}{p_{2,i}} \gamma_i x_{1,i,t}^\beta_{1,i} \beta_{2,i} x_{2,i,t}^{\beta_{2,i}-1}$$ (14)

where p_1 and p_2 are unit prices of the inputs x_1 and x_2. Separate multipliers are introduced to allow for the inputs to be constrained by different budget allocations. Equation (14) implies that any area with positive allocation must have γ_i > 0, and positive allocations in more than one area (i) imply 0 < (β_{1,i} + β_{2,i}) < 1 for all areas.

Assume further than institutional constraints require that total allocation of the inputs to each area before t=1. In practice, this occurs when budgets are allocated on a yearly basis, but expenditures are deployed on a monthly basis. Possible deviations from the optimal allocation of

$$\sum_{t=1}^{T} x_{1,i,t} \text{ and } \sum_{t=1}^{T} x_{2,i,t} \text{ over time are ignored. Thus we have:}$$

$$\sum_{t=1}^{T} y_{i,t} = T \beta_{0,i} + \gamma_i T \left( \frac{\sum_{t=1}^{T} x_{1,i,t}}{T} \right)^{\beta_{1,i}} \left( \frac{\sum_{t=1}^{T} x_{2,i,t}}{T} \right)^{\beta_{2,i}} + \sum_{t=1}^{T} \epsilon_{i,t}$$ (15)
and the optimal allocations \( \left( \sum_{t=1}^{T} x_{i,j,t} \right)^* \) and \( \left( \sum_{t=1}^{T} x_{2,i,t} \right)^* \) are just \( T \times (x_{i,j,t})^* \) and \( T \times (x_{2,i,t})^* \) where \( (x_{1,i,t})^* \) and \( (x_{2,i,t})^* \) are obtained from (14).

Error terms are added to equation (14) to allow for suboptimal allocation:

\[
\lambda_1 = \frac{dy_{i,t}}{p_{i,t} dx_{i,t}} = \frac{1}{p_{i,t}} \gamma_i \beta_{1,i} x_{i,t}^{\beta_{1,i}-1} x_{2,i,t}^{\beta_{2,i}} \exp(\xi_{1,i}) \\
\lambda_2 = \frac{dy_{i,t}}{p_{2,t} dx_{2,i,t}} = \frac{1}{p_{2,t}} \gamma_i \lambda_{i,j} \beta_{1,i} x_{i,t}^{\beta_{1,i}-1} \exp(\xi_{2,i})
\]  

Taking logs and rearranging we obtain

\[
\ln x_{i,t} = \left[ \begin{array}{c}
\ln \left( x_{i,j,t} \right) \\
\ln \left( x_{2,i,t} \right)
\end{array} \right] = \left[ \begin{array}{cc}
\beta_{1,i} - 1 & \beta_{2,i} \\
\beta_{1,i} & \beta_{2,i} - 1
\end{array} \right]^{-1} \left[ \begin{array}{c}
\ln \lambda_1 - \ln \gamma_i - \ln \beta_{1,i} + \ln p_{i,t} + \xi_{1,i} \\
\ln \lambda_2 - \ln \gamma_i - \ln \beta_{2,i} + \ln p_{2,t} + \xi_{2,i}
\end{array} \right]
\]  

with \( \ln \left( \sum_{t=1}^{T} x_{i,t} \right)^* = \ln x_{i,t} + \ln T \).

After \( \left( \sum_{t=1}^{T} x_{i,j,t} \right)^* \) and \( \left( \sum_{t=1}^{T} x_{2,i,t} \right)^* \) are thus determined for all areas, we assume that control is handed over to the local management. Deployment of the total input over time may depart from the implied optimal deployment over time for a multitude of reasons that are beyond consideration and summarized by \( \eta_i \). The set \( \{ \eta_i \} \) is assumed to be orthogonal to all parameters of interest.

The data generating mechanism described above can be summarized by the directed acyclic graph (DAG) displayed in figure 1 (see Liu, Otter and Allenby 2007, Rossi, Allenby and McCulloch 2005). Three geographic areas are depicted in the DAG, indexed by \( i = \{1, i, N\} \). The parameters \( \theta_i = \{ \beta, \gamma_i \} \) are unit-level parameters that influence the allocation made to the unit, and the response to the allocation. Also affecting the allocation decision are the multiplier (\( \lambda_i \)), price (\( p_i \)), and allocative error (\( \zeta \)). Additional error (\( \eta \)), which is not modeled, takes the aggregate (e.g., yearly) allocation and
distributes it on a finer scale (e.g., monthly). Outputs, \{y_i\}, are influenced by \(\theta_i\) directly, and indirectly through the allocation \(\sum_{t=1}^T x_{i,t}\).

If sufficient data are available such that pooling is not necessary, then analysis for this model that proceeds on a unit-by-unit basis is not prone to the biasing effects of endogeneity. The reason is that, at the unit-level, the allocation \(\sum_{t=1}^T x_{i,t}\) is constant. Estimates of \(\theta_i\) can be obtained from variation of the inputs (x) and output (y) variables over time using a conditional demand model. However, when pooling is present in the model, either by aggregating the data or through a random-effect specification, the allocation of expenditures vary by unit according to \(\theta_i\) and we show below that the supply-side specification becomes important.

The parameters \(\tau\) and \(\Sigma\) are unobserved hyper-parameters describing the population of areas in terms of their productivity and the amount of error in the allocation decisions, respectively. We note that \(\lambda\) and \{\(p_i\)\} may or may not be observed – we develop inference procedures assuming they are not observed at the cost of a somewhat ambiguous interpretation of \{\(\zeta_i\)\}. With \{\(p_i\)\} unobserved, \{\(\zeta_i\)\} may reflect price variations across areas known by the decision maker but unobserved by the analyst and/or pricing errors by the decision maker. We comment further on this aspect of the analysis below.

The DAG shows that \(\tau\), after integrating out \{\(\theta_i\)\}, depends on the outputs \{\(y_i\)\} and inputs \(\left\{\sum_{t=1}^T x_{i,t}\right\}\). Thus inference for \(\tau\) based on demand data alone is generally inconsistent. However, when the number of time periods \(T\) is large, \(\sigma^2\) is small, and there is sufficient variation in the inputs, then one can obtain accurate estimates of \{\(\theta_i\)\} from the demand-side equations exclusively and
inferences for $\tau$ are consistent. The reason is that $\tau$ is independent of both $\{y_i\}$ and $\left\{\sum_{t=1}^{T} x_{i,t}\right\}$ given $\{\theta_i\}$. However, this approach to estimation can be inaccurate when the number of time periods $T$ is small, $\sigma^2$ is relatively large, or allocation over time is already optimal and thus constant. In general, the demand-side likelihood leaves considerable uncertainty about $\theta_i$ such that data pooling is desirable, and as a result a supply-side specification is necessary.

We created 30 replications of populations of areas each consisting of 200 areas conditional on $\tau$. It is assumed that managers base their yearly budget allocations on knowledge about the responsiveness of individual areas in a population $\{\theta_i\}$, the attractiveness of other investment options, $\lambda$, and input prices, $\{p_i\}$, which we take to be constant across areas in the simulation. After the yearly budgets are assigned to each geographic area, area management deploys the budget over a period of twelve months ($T = 12$). Reasons for departing from the optimal, constant allocation over time, $\{\eta_i\}$, are assumed exogenous to the system. In the simulation we use $\eta_i \sim \text{Dirichlet}(1, \ldots, 1_T)$ to divide the yearly budget into monthly budgets that add to the yearly budget. Note that the choice of parameters in the Dirichlet distribution leads to large month to month variation in the deployed budgets. Given the allocations $x_{i,t}$, we then generated monthly outputs according to equation (13).

We assume that the local area management is ignorant about the causes underlying $\{\varepsilon_i\}$ such that $\{\eta_i\}$ and $\{\varepsilon_i\}$ are independent. The $\{\eta_i\}$ reflect local problems that stand in the way of optimal allocation over time.

Appendix A provides Bayesian MCMC algorithms for estimating the model described in figure 1. Table 1 reports results for selected parameters. We find that estimates based on the conditional demand model to be off, with response coefficients ($\ln \beta$) estimated to be too large and the multiplicative intercept to be too small. Our finding of an increase in the estimated responsiveness to the input variables is similar the bias encountered in simple pooling of data in a
standard regression model, discussed above in the introductory section. Finally, we also find incorrect inferences for elements of the random-effects covariance matrix. The results, in general, point to the importance of incorporating the supply-side into the model.

The bottom portion of table 1 reports the expected gains from optimally allocating input $x_1$ while holding input $x_2$ fixed at its observed value. Two improvements are investigated – one based on the posterior distribution from the demand-side model, and the other based on the posterior that includes the supply-side model. Recall that the actual inputs in the simulation deviate from optimality because of the supply-side error $\zeta$. We define gain as the expected increase in output from reallocating the inputs, subject to the constraint that the total input level is held fixed:

$$
E\left[ \sum_{i,t} y_{i,t}^\text{opt} | x_{1,i,t}^\text{opt}, x_{2,i,t}, \pi(\theta | \cdot) \right] - E\left[ \sum_{i,t} y_{i,t} | x_{1,i,t}, x_{2,i,t}, \pi(\theta | \cdot) \right] \quad \text{s.t.} \quad \sum_{i,t} x_{1,i,t}^\text{opt} = \sum_{i,t} x_{1,i,t} \quad (18)
$$

where the expectation is taken with respect to the posterior distribution $\pi(\theta | \cdot)$ and the 30 simulated data sets. We find that gains based on the demand model are too optimistic, with values four times larger than those obtained when the supply-side model is also used in parameter estimation – i.e., 805 versus 171. We find that estimates using the supply-side model agree with the gains obtained when using the true parameter values instead of the posterior distribution (i.e., 221). Finally, we examine performance of identified optimal allocation above using the true parameter values, not the posterior distribution $\pi(\theta | \cdot)$ in equation (18). We find that the optimal allocation of $x_1$ from the demand-side only model results in a loss of output when evaluated with the true parameters (-378), while the optimal allocation based on the demand- and supply-side model results in a gain of 151. These results indicate that ignoring the supply side results in incorrect inferences of model parameters and sub-optimal allocation of inputs.
4. Evaluating Marketing Expenditures

*Data and Models*

We illustrate our model with data from a services organization operating in 21 regions along the eastern seaboard of the United States. The firm maintains multiple branch offices in each of the regions, and makes yearly promotional allocations that are then spent over the course of the year. One of the outputs generated by these inputs is the net number of new customers per month, the dependent variable in our analysis. There are 294 observations available for analysis (14 months of data for each of 21 regions).

Promotional expenditures \( (x_{1i}) \) are plotted against the number of branches \( (x_{2i}) \) and the region’s population in logarithmic form in figure 2. The points in each of the plots fall along the 45 degree line, corresponding to a unit slope, or a proportional allocation rule. That is, it appears that promotional expenditure and the number of branch outlets varies in direct proportion to the population in each region. The question to be addressed is whether this allocation is optimal given the efficiency of the geographic regions.

\[ \text{Figure 2} \]

It is useful to think of the effect of marketing expenditures in per capita terms. Promotional expenditures and the number of branch outlets have diminishing marginal effects that should depend on the population size of the region, with allocations made in a smaller market areas satiating more quickly. We therefore standardize the variables in the analysis per 1000 population, and view each (standardized) geographic region as providing exchangeable information for
estimating model parameters. This standardization allows us to test whether the geographic regions act as homogeneous production units by testing if model parameters are equal across regions.

Figure 3 displays the distribution of new customers per month per 1000 population for each of the regions, the dependent variable in the analysis reported below. Despite this standardization, we find variation in the distributions, suggesting that the residual variance in the production functions should not be constrained across regions.

== Figure 3 ==

We specify and test alternative versions of a multiplicative and an additive production function. The multiplicative model is similar to the model studied in the simulation study (equation 13), and the additive model assumes the inputs do not interact:

\[ y_{i,t} = \beta_0 + \gamma_1 x_{1,i,t-1} + \gamma_2 x_{2,i} + \epsilon_{i,t} \]  \hspace{1cm} (19)

\[ y_{i,t} = \beta_0 + \gamma_1 x_{1,i,t-1} + \gamma_2 x_{2,i} + \epsilon_{i,t} \]  \hspace{1cm} (20)

where \( y_{i,t} \) denotes the number of new customers produced by region \( i \) in month \( t \), \( x_{1,i,t-1} \) is lagged promotional expenditure in region \( i \), and \( x_{2,i} \) denotes the number of branch offices in the region that does not vary over the span of our data. Versions of these models specify fixed and random-effects for the coefficients, and we investigate performance of the models with and without the supply-side specification.

Figure 4 displays the input variables in the analysis. The top portion of the figure plots promotional expenditures per 1000 population for each of the 21 regions in the analysis. The plot shows that monthly promotional expenditure is often zero, but has sufficient variation to facilitate estimation of the input coefficient \( \beta_1 \). The bottom portion of the figure displays the number of branch outlets per 1000 population in each region. The number of branches in a region does not vary over the length of the data, and therefore are not plotted against time.
We investigate separate budget allocations for total promotional expenditures and the number of branch offices. The data provided by the firm does not include input prices for these two variables, and we therefore cannot estimate a common budget constraint parameter (i.e., Lagrangian multiplier) \( \lambda \) used to form the auxiliary function. Instead we assume that a separate budget allocation process for these inputs, which is reasonable, for example, when branch offices are considered a capital expense item, and promotional expenditures are considered period expenses. In addition, we assume that the allocation decision for promotional expenditures are made once for all time periods, and that the deployment of these resources is governed by other factors such as the timing of media purchases and extraneous influences independent of the model parameters.

**Results**

In-sample and predictive fits for alternative specifications of the multiplicative (equation 19) and additive (equation 20) models are displayed in table 2. The top portion of the table displays results for the multiplicative model, and the bottom portion of the table displays results for the additive model. The first column of the table indicates parameters that are constrained to be the same across geographic regions. The second column indicates the presence and manner in which the supply-side model is included. A "no" indicates that only the demand-side model is used to obtain parameter estimates, "yes" indicates the presence of the supply-side equations, and "\( \Sigma \) large" indicates that the error variance in the supply side specification is a priori set to be large. Fixing \( \Sigma \) to be large results in ordinal parameter constraints from the supply-side specification because of the Jacobian in, for example, equation (12), but little other information. Comparison of "yes" to "\( \Sigma \) large" allows assessment of optimal allocation behavior by the firm.
The third column reports a measure of in-sample fit (the log marginal density), and the last two columns of the table report predictive results for each of the models holding out the last observation. As usual, the posterior distribution from the calibration data is used to predict the held-out observation. Depending on the model this posterior distribution is informed / constrained by the supply side to various degrees (see also Manchanda, Rossi and Chintagunta (2004)). The advantage of this method is that it facilitates comparison across all models, including those without the supply-side specification. If the supply-side model is mis-specified, it will degrade predictions relative to a corresponding model with a less informative or unspecified supply-side. Two predictive measures are displayed – the log predictive density and the mean squared error.

The multiplicative model with a fixed-effect multiplicative constant ($\gamma$ fixed) is found to be the best fitting model. It has the highest log marginal density of all models with a supply-side present, and results in the best predictive mean squared error across any model. It also has the second highest log predictive density, worse only to the demand-side only model with input coefficients $\beta_1$ and $\beta_2$ specified as fixed-effects. In general, we find the multiplicative models fit better than the additive models. We also find that fits of the best model improve with the addition of supply-side information, including both the ordinal constraints ($\Sigma$ fixed) and the full specification.

== Table 2 ==

Table 3 reports parameter estimates for the best-fitting model – the multiplicative model (equation 19) with the multiplier $\gamma$ specified as a fixed-effect. All other coefficients in the model are specified as random-effects. The top portion of table 3 displays posterior means and standard deviations of the product function hyper-parameters and the fixed-effect. The bottom portion of the table displays the allocation parameters unique to the supply-side equations.

We note that the Jacobian associated with the supply-side likelihood allows us to distinguish fixed- versus random-effect specification, even when there is no temporal variation in an input (e.g.,
With $\beta$, fixed across geographies, the Jacobian imposes the restriction $0 < \beta_1 + \beta_2 < 1.0$, while for the random-effects model the restriction is $0 < \beta_{1i} + \beta_{2i} < 1.0$. Thus, the fixed-effect model places greater constraints on the response coefficients for the first input, $\beta_{1i}$, for non-zero optimal allocations to exist.

Estimates of the effect of promotional expenditure and branch offices are both small, with the mean promotional expenditure coefficient equal to 0.01. The branch office coefficient is four times larger. The small magnitude of the expenditure coefficient essentially reduces the effect of promotions, $x_{ii}^{11}$, to a step function that equals zero when $x=0$ and equals one when $x>0$. The implication is that large promotional expenditures have the same effect as small promotional expenditures in the range of expenditures investigated.

When promotional expenditures are positive, their effect depends on the number of branch outlets (per 1000 population) in the region. However, when promotional expenditures are zero, the branch outlets do not help to generate new customers. Thus, the results indicate a strong interactive effect that is consistent with an effects hierarchy as discussed in equation (1). In the data examined, promotional expenditures are equal to zero 76% of the time. The model suggests that gains may be available by increasing the frequency of low dollar promotions by reducing the magnitude of high dollar promotions so that promotions occur in every period. We investigate this issue below.

== Table 3 ==

Figure 5 displays posterior estimates of error variances ($\sigma_i^2$) for each of the 21 regions in the analysis. The posterior distributions differ across regions, providing support for the modeling assumption of region-specific variance.

== Figure 5 ==
5. Discussion

Evaluating the effectiveness of marketing expenditures requires inferences about a system of allocation and response under constraints of limited information. Analysts are interested in alternative formulations of allocation (i.e., the supply-side), including the special case where $p(x)$ is left unspecified. Leaving $p(x)$ unspecified is desirable because there is usually no theory available that guides the choice of the supply-side likelihood independent of the demand equations and its parameters. Bayes theorem provides an elegant solution to this problem. The key is to notice that $x$ can be used to update the prior distribution of the parameters in the demand equation $\pi(\theta)$ to form $\pi(\theta|x)$. Once this normalized (i.e., proper) density is obtained we can meaningfully compare the marginal likelihoods of the demand data under two different prior distributions, where one prior distribution is informed by the supply side.

A problem with this method is that the supply-side likelihood contribution to the posterior distribution $\pi(\theta|x)$ usually does not uniquely identify $\theta$. In our examples above, the $x$ variables are assumed to have been set as controls for the demand-side variable, $y$, and the decision maker's understanding of the mapping from $x$ to $y$ underlying $\pi(x|\theta)$ only identifies subspaces of $x$ as more likely. The topology of those subspaces usually does not have a (unique) maximum. However, with somewhat informative $y$, it is often straightforward to sample from the posterior distribution using a calibration dataset, and compute predictive densities for holdout observations. This can also be done with the supply-side left unspecified.

In our model we have additional parameters in the supply side likelihood $\pi(x|\theta, \cdot)$, summarized by a dot, that specify how close the supply side is to the optimum implied by the demand side. In this situation, informative priors on these parameters may be used to assess
whether fitting the supply side results in a poor fit to the demand data. This is important, since the joint fit to the demand and the supply side data may be masking that adding the supply side negatively affects the fit of the demand side which should call the model formulation into question. In our example, an a priori fixed and very large variance for the error terms on the supply side leaves the supply side likelihood essentially flat such that the posterior is (almost entirely) driven by the demand side likelihood. If such a specification results in a much improved fit to the demand data one should question the idea that managers pick the x variables based on an understanding of the parameters in the demand equation.

A feature of specifying a supply side that is essentially flat with regard to the size of the allocation errors is that it still ensures a posterior that is consistent with the existence of a 'well-behaved' global optimum. The implied optimal action can be reconciled with basic features of the observed data, such as investments of a company in more than one area. This is, of course, not the case of \( p(x) \) if left unspecified. In our analysis, we find support for the supply-side implications of our best fitting model.

The small magnitude of the response coefficients \( \beta_1 \) and \( \beta_2 \) imply that it would be advantageous to the firm to spend promotional dollars in each period, rather than the pattern observed in figure 4 where promotion levels are at zero 76% of the time. While our model assumes that the monthly deployment of promotional dollars is exogenous to our system of study, we can calculate the gain from "filling-in" the zero expenditures by reallocating the non-zero expenditures within each branch. We estimate this action to lead to a two-fold increase in the number of new customers per capita.

Our model can be used to make predictions of optimal allocations across regions. Figure 6 plots the current versus the optimal promotional spending allocation for the regions. The magnitude of the reallocation ranges up to 20%, and we find the pattern of reallocation unrelated to
the region population, nor the per capita number of branches. The location of the posterior
distribution of response coefficients near zero (see table 3), when entered in logarithmic form in the
conditional supply-side equation (17), results in large deviations between current and optimal
spending. Unfortunately, these same small coefficients essentially lead to no expected gain in
expected demand when monthly changes in spending are constrained to be proportional to
observed spending so that the zero promotional periods are preserved. Alternative allocations
based on different regional costs results in the same finding.

== Figure 6 ==

6. Conclusion

This paper introduces a method of assessing the effectiveness of marketing expenditures. It
employs a direct estimation strategy that does not rely on a dual cost representation, and is
sufficiently general to be used with the variety of non-standard production functions encountered in
marketing applications. Our method can address the efficiency of production units (regions in our
analysis) to convert inputs to outputs given the existing allocations (i.e., technical efficiency), as well
as the optimality of current allocations (i.e, allocative efficiency). Technical efficiency is viewed in
terms of heterogeneous model coefficients, and allocative efficiency is modeled with a series of
conditional first order conditions. We show how our framework can be used to make inferences
about the allocation behavior of the firm that may be different from the optimal behavior implied by
the demand equation.

We demonstrate through a simulation experiment that ignoring the expenditure allocation
rule results in incorrect inference whenever allocations vary across units of analysis. The simulation
study also demonstrates that our proposed solution leads to valid inferences. In application to data
from a service provider operating in multiple geographic regions, we find evidence that the regions
are differentially efficient at converting the inputs promotional spending and branch offices into new customers. While we find that the inputs are sub-optimally allocated across regions, the small value of the estimated response coefficients lead to essentially no gain in the number of new customers. Output gains from optimal allocation are shown to be much larger in the simulation study.

Our analysis demonstrates the flexibility of our method in assessing efficiencies despite not having access to input prices or other variables that are influential in generating new business. It is doubtful that analysts in marketing ever have a complete set of variables for analysis, and it is critical that analysis can proceed with functional forms that potentially accommodate their absence through additive intercepts and error terms.

The framework and analysis presented in this paper can be extended in a number of ways. First, our investigation involved aspects of promotion and distribution, but did not involve product-related expenditures. Data were also not available on promotional quality, media content, or differing regional costs of branch outlets. The assessment of input prices is often difficult, being dependent on cost drivers and cost allocation rules used by firms, and the development of input price estimators is an interesting extension of our work. Finally, we feel the development and comparison of alternative production functions for marketing would constitute important contributions to the literature.
Appendix A

Estimation Algorithms

Estimation is based on MCMC methods. We outline the sampler for the multiplicative model. The
sampler for the additive model follows the same general outline. Here we employ a centered
parameterization where the mean of the random effects is independent of the data, conditional on
the random effects. When we specify one or more effects as fixed, i.e. homogenous, we employ a
non-centered parameterization, where the fixed effects (and the mean of the random effects) are
proposed from the posterior of the demand side and accepted or rejected in a Metropolis step that
involves the joint supply side likelihood of all areas.

The joint posterior density of all model parameters is factored as follows:

\[ p \left( \{ \beta_i, \gamma_i, \kappa_i \}, \sigma^2, \alpha, Q, \ln \lambda, \Sigma \left| \{ y_i, X_i \} \right. \right) \propto \]

\[ \prod_{i=1}^{N} \prod_{t=1}^{T} p \left( y_{i,t} \left| x_{i,t}, \beta_i, \gamma_i, \sigma^2 \right. \right) \times \]

\[ \prod_{i=1}^{N} p \left( \ln \sum_{t=1}^{T} x_{i,t} - \ln T \left| \beta_{1,i}, \beta_{2,i}, \ln \gamma_i, \ln \lambda, \Sigma \right. \right) \times \]

\[ \prod_{i=1}^{N} p \left( \beta_i, \gamma_i \left| a, Q \right. \right) \times \]

\[ p \left( \ln \lambda \right) p \left( \Sigma \right) p \left( a \right) p \left( Q \right) p \left( \sigma^2 \right) p \left( \{ \kappa_i \} \right) \]

\[ p \left( \beta_i, \gamma_i \left| a, Q \right. \right) \text{ is assumed to be multivariate normal and we use the usual conditionally conjugate,} \]

weakly informative prior distributions. Parameters \{ \kappa_i \} account for cross sectional heterogeneity in
the demand observation error variance, i.e. \( \sigma^2_{\varepsilon_i} = \frac{\sigma^2_{\varepsilon_i}}{\kappa_i} \) with \( p \left( \{ \kappa_i \} \right) = \prod_{i=1}^{N} \text{Gamma} \left( \kappa_i | 2, 2 \right). \)

The MCMC sampler cycles through the following blocks
1. \( p(a|\{\beta_i, \gamma_i\}, Q) \)

2. \( p\left( \ln \lambda \left| \sum_{i=1}^{T} x_{i,i}, \{\beta_{1,i}, \beta_{2,i}, \ln \gamma_i\}, \Sigma \right. \right) \)

3. \( p\left( \Sigma \left| \sum_{i=1}^{T} x_{i,i}, \{\beta_{1,i}, \beta_{2,i}, \ln \gamma_i\}, \ln \lambda \right. \right) \)

\[ p\left( \{\beta_{0,i}, \gamma_i\}^\ast \right) \propto \prod_{i=1}^{N} p\left( \beta_{0,i}, \gamma_i \mid y_{i,i}, x_{i,i}, \beta_{1,i}, \beta_{2,i}, \sigma^2_x, \kappa_i, a, Q \right) \times \]

4. \( p\left( \ln \sum_{i=1}^{T} x_{i,i} - \ln T \left| \beta_{1,i}, \beta_{2,i}, \ln \gamma_i, \ln \lambda, \Sigma \right. \right) \)

\[ p\left( \{\beta_{1,i}, \beta_{2,i}\}^\ast \right) \propto \prod_{i=1}^{N} p\left( y_{i,i} \mid x_{i,i}, \beta_{i}, \gamma_i, \sigma^2_x, \kappa_i \right) \times p\left( \beta, \gamma_i \mid a, Q \right) \times \]

5. \( p\left( \ln \sum_{i=1}^{T} x_{i,i} - \ln T \left| \beta_{1,i}, \beta_{2,i}, \ln \gamma_i, \ln \lambda, \Sigma \right. \right) \)

6. \( p\left( \{\kappa_i\} \mid \{y_{i,i}, x_{i,i}, \beta_i, \gamma_i, \sigma^2_x\} \right) \)

7. \( p\left( \sigma^2_x \mid \{y_{i,i}, x_{i,i}, \beta_i, \gamma_i, \kappa_i\} \right) \)

8. \( p(Q|\{\beta_i, \gamma_i\}, a) \)

Blocks 1 to 3 as well as 6, 7 and 8 are standard Gibbs samplers.

In block 4 we take advantage of the fact that equation (13) is a simple normal, linear regression conditional on \( \beta_{1,i} \) and \( \beta_{2,i} \). Thus we employ \( p\left( \beta_{0,i}, \gamma_i \mid y_{i,i}, x_{i,i}, \beta_{1,i}, \beta_{2,i}, \sigma^2_x, \kappa_i, a, Q \right) \) as a proposal density and are left with only \( p\left( \ln \sum_{i=1}^{T} x_{i,i} - \ln T \left| \beta_{1,i}, \beta_{2,i}, \ln \gamma_i, \ln \lambda, \Sigma \right. \right) \) in the Metropolis ratio. This latter density is obtained by rearranging equation (17) and employing change of variable calculus. Note that the Jacobian matrix in this case is equal to the identity.
The Jacobian matrix involved in the move from proposing $\gamma_i$ to evaluating $\ln \gamma_i$ in the metropolis step cancels from the metropolis ratio.

In block 5,\[ p \left( \{ \beta_{1,i}, \beta_{2,i} \} \mid y_i \right) \propto \prod_{i=1}^{N} p \left( y_{i,t} \mid x_{i,t}, \beta_{1,i}, \gamma_i, \sigma_{\epsilon}, \kappa_i \right) p \left( \beta_{1,i}, \gamma_i \mid \alpha, Q \right), \] i.e. the (conditional) demand side contribution to the posterior, cannot be sampled from directly as $\beta_{1,i}$ and $\beta_{2,i}$ appear in the exponent of equation 13. We therefore employ a random walk Metropolis sampler in block 5. The Jacobian, i.e. the absolute value of the determinant of the Jacobian matrix, from the supply side is\[ \left| \begin{bmatrix} \beta_{1,i} - 1 & \beta_{2,i} \\ \beta_{1,i} & \beta_{2,i} - 1 \end{bmatrix} \right| = \left| \begin{bmatrix} \beta_{1,i} - 1 \\ \beta_{2,i} - 1 \end{bmatrix} \right| = \left| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right| = 1. \]

For positive $\beta_{1,i}$ and $\beta_{2,i}$, the Jacobian effectively enforces diminishing marginal returns as it results in a likelihood of zero for all $(\beta_{1,i} + \beta_{2,i}) = 1$.

**Optimal Actions**

We describe how to obtain the optimal action for the multiplicative model. The same general principles apply to any other model of demand that implies an interior globally optimal allocation.

If the response parameters were known, the optimal action $x_{i,t}$ simply solves the following equation for $\zeta = 0$.

\[
\begin{bmatrix}
\beta_{1,i} - 1 & \beta_{2,i} \\
\beta_{1,i} & \beta_{2,i} - 1
\end{bmatrix}
\begin{bmatrix}
\ln x_{i,t} \\
\ln x_{i,t}
\end{bmatrix}
= \begin{bmatrix}
\ln \lambda_i - \ln \gamma_i - \ln \beta_{1,i} + \ln p_1 \\
\ln \lambda_i - \ln \gamma_i - \ln \beta_{2,i} + \ln p_2
\end{bmatrix}
= \begin{bmatrix}
\zeta_{1,i} \\
\zeta_{2,i}
\end{bmatrix}
\]
In our application we lack the price information needed to optimally balance the two inputs $x_{1,i,t}$ and $x_{2,i,t}$ within areas. Thus, we investigate the conditionally optimal allocation of the first input across areas assuming optimal allocation across time. Note that we implicitly assume that the price of the first input is constant across areas.

Thus, we solve the equation \( (\beta_{i,j} - 1) \ln x_{1,i,t} + \beta_{2,j} \ln x_{2,j} - \ln \lambda_i + \ln \gamma_i + \ln \beta_{i,j} = 0 \) for the first input:

\[
\ln x_{1,i,t} = \frac{-\beta_{2,j} \ln x_{2,j} + \ln \lambda_i - \ln \gamma_i - \ln \beta_{i,j}}{\beta_{i,j} - 1} = \frac{\beta_{2,j} \ln x_{2,j} - \ln \lambda_i + \ln \gamma_i + \ln \beta_{i,j}}{1 - \beta_{i,j}}
\]

Note that for given \( \{x_{2,j}\} \), Lagrangian multiplier \( \lambda_i \), and response parameters this system of equations implies that the sum of the log of the optimal \( \{x_{1,i}\} \) is equal to the expected sum of the log of the observed \( \{x_{1,i}\} \), where the expectation is over the supply side errors, \( \{\zeta_i\} \). For a cost-neutral optimal policy that does not depend on the existence of alternative investment options and their unobserved productivity, we require that the sum of the optimal \( \{x_{1,i}\} \) is equal to the sum of the observed \( \{x_{1,i}\} \). We accomplish this by solving the system of N equations (one equation per area) for the Lagrangian multiplier that satisfies this constraint.

For the optimal action to be a Bayes rule, we have to take full account of any uncertainty remaining in our assessment of the productivity of the different areas. We accomplish this by solving for the optimal, cost-neutral action at each MCMC draw from the posterior. The optimal Bayes action then is the average of the optimal actions obtained at each draw.
References


Figure 1
Direct Acyclic Graph for Simulation Study
Figure 2
Regional Dispersion of Promotional Expenditures, Branch Outlets and Population
Figure 3
Regional Distributions of New Customers per Month per 1000 Population
Figure 4
Input Variables

Promotional Expenditure per 1000 Population

Branch Outlets per 1000 Population
Figure 5
Posterior Distributions of Error Variance ($\sigma_i^2$)
Figure 6
Observed and Optimal Promotional Expenditures Conditional on Branch Outlets
Table 1
Simulation Study Parameter Estimates
(Standard Error of Posterior Means across 30 Simulated Data Sets)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Value</th>
<th>Demand-Side Model</th>
<th>Demand and Supply-Side Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand Side Mean of Random-Effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(\beta_1)$</td>
<td>-1.50</td>
<td>-1.38 (0.01)</td>
<td>-1.49 (0.01)</td>
</tr>
<tr>
<td>$\ln(\beta_2)$</td>
<td>-1.50</td>
<td>-1.38 (0.01)</td>
<td>-1.49 (0.01)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>3.50</td>
<td>2.94 (0.06)</td>
<td>3.46 (0.07)</td>
</tr>
<tr>
<td>Variance of Random-Effects:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(\beta_1)$</td>
<td>0.080</td>
<td>0.062 (0.002)</td>
<td>0.081 (0.002)</td>
</tr>
<tr>
<td>$\ln(\beta_2)$</td>
<td>0.080</td>
<td>0.071 (0.002)</td>
<td>0.081 (0.002)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.950</td>
<td>0.753 (0.027)</td>
<td>0.874 (0.039)</td>
</tr>
<tr>
<td>Supply Side Lagrange Multiplier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\ln(\lambda_1)$</td>
<td>-2.30</td>
<td>NA</td>
<td>-2.30 (0.02)</td>
</tr>
<tr>
<td>$\ln(\lambda_2)$</td>
<td>-2.30</td>
<td>NA</td>
<td>-2.32 (0.04)</td>
</tr>
<tr>
<td>Variance of Allocation Error</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Sigma(1,1)$</td>
<td>0.01</td>
<td>NA</td>
<td>0.011 (0.001)</td>
</tr>
<tr>
<td>$\Sigma(2,2)$</td>
<td>0.30</td>
<td>NA</td>
<td>0.299 (0.006)</td>
</tr>
<tr>
<td>Gains from Optimal Actions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gain $</td>
<td>x_1^{\text{opt}}, \pi(\theta</td>
<td>\cdot)^T$</td>
<td>221 (26)</td>
</tr>
<tr>
<td>Gain $</td>
<td>x_1^{\text{opt}}, \theta^{\text{true}}$</td>
<td>221 (26)</td>
<td>-378 (33.8)</td>
</tr>
</tbody>
</table>

$^1$See equation (18)
Table 2
Model Fit

| fixed effects | Supply Side | ln p(y, x | M)  | LPD  | MSE  |
|---------------|-------------|-------------|------|------|
| Multiplicative Model: $y_{i,t} = \beta_{0,t} + \gamma_{0}x_{i,t-1}^{\beta_{0},} + \gamma_{1}x_{2,t}^{\beta_{1},} + \epsilon_{i,t}$ |
| $\gamma$ | no | NA | 15.40 | 2.24 |
| $\beta_{2}$ | no | NA | 15.65 | 2.31 |
| $\beta_{1}, \beta_{2}$ | no | NA | 16.21 | 2.29 |
| $\gamma, \beta_{2}$ | no | NA | 14.78 | 2.16 |
| $\gamma$ | yes | | 360.40 | 15.73 | 2.14 |
| $\beta_{2}$ | yes | | 331.60 | 12.95 | 2.55 |
| $\beta_{1}, \beta_{2}$ | yes | | 318.02 | 12.64 | 2.47 |
| $\gamma, \beta_{2}$ | yes | | 326.82 | 15.14 | 2.14 |
| $\gamma$ | $\Sigma$ large | | 119.62 | 15.21 | 2.19 |
| $\beta_{2}$ | $\Sigma$ large | | 123.91 | 14.79 | 2.61 |
| $\beta_{1}, \beta_{2}$ | $\Sigma$ large | | 124.30 | 14.43 | 2.53 |
| $\gamma, \beta_{2}$ | $\Sigma$ large | | 118.15 | 15.03 | 2.16 |

| fixed effects | Supply Side | ln p(y, x | M)  | LPD  | MSE  |
|---------------|-------------|-------------|------|------|
| Additive Model: $y_{i,t} = \beta_{0,t} + \gamma_{1}x_{i,t}^{\beta_{0},} + \gamma_{2}x_{2,t}^{\beta_{1},} + \epsilon_{i,t}$ |
| $\beta_{0}, \gamma_{2}$ | no | NA | 14.72 | 2.62 |
| $\beta_{0}, \beta_{2}$ | no | NA | 14.80 | 2.57 |
| $\gamma_{2}, \beta_{2}$ | no | NA | 14.63 | 2.59 |
| $\beta_{0}, \gamma_{2}$ | yes | | 294.66 | 11.17 | 2.85 |
| $\beta_{0}, \beta_{2}$ | yes | | 307.69 | 12.86 | 2.68 |
| $\gamma_{2}, \beta_{2}$ | yes | | 328.66 | 13.94 | 2.78 |
| $\beta_{0}, \gamma_{2}$ | $\Sigma$ large | | 102.81 | 14.00 | 2.61 |
| $\beta_{0}, \beta_{2}$ | $\Sigma$ large | | 124.62 | 14.32 | 2.61 |
| $\gamma_{2}, \beta_{2}$ | $\Sigma$ large | | 122.94 | 15.28 | 2.62 |

1 All other effects are specified as random. 2 When $\beta_{0}$ is fixed it is a priori fixed to zero. 3 Harmonic Mean Estimator (Newton and Raftery 1994). 4 Log predictive density and mean squared prediction error for the held out demand observations.
### Table 3
Parameter Estimates for Preferred Model ($\gamma$ Fixed)
Posterior Mean (Posterior Standard Deviation)

<table>
<thead>
<tr>
<th>Production Function Parameters</th>
<th>Intercept</th>
<th>$\bar{\beta}_0$</th>
<th>$V_{\beta_0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.189</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.079)</td>
<td>(0.039)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Promotional Expenditure</th>
<th>$\log \bar{\beta}_1$</th>
<th>$V_{\log \beta_1}$</th>
<th>$\bar{\beta}_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-5.064</td>
<td>0.198</td>
<td>0.0091</td>
</tr>
<tr>
<td></td>
<td>(0.748)</td>
<td>(0.064)</td>
<td>(0.0087)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Branch Offices</th>
<th>$\log \bar{\beta}_2$</th>
<th>$V_{\log \beta_2}$</th>
<th>$\bar{\beta}_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-3.680</td>
<td>0.156</td>
<td>0.0416</td>
</tr>
<tr>
<td></td>
<td>(1.095)</td>
<td>(0.050)</td>
<td>(0.0345)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Multiplier</th>
<th>$\gamma$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.326</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Allocation Parameters</th>
<th>Intercepts:</th>
<th>$\log \lambda_i$</th>
<th>$\log \lambda_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-7.672</td>
<td>-3.658</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.773)</td>
<td>(1.157)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Allocative Error Variance:</th>
<th>$\Sigma^{(1,1)}_\zeta$</th>
<th>$\Sigma^{(1,2)}_\zeta$</th>
<th>$\Sigma^{(2,2)}_\zeta$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0015</td>
<td>-0.0001</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>(0.0033)</td>
<td>(0.0019)</td>
<td>(0.0031)</td>
</tr>
</tbody>
</table>