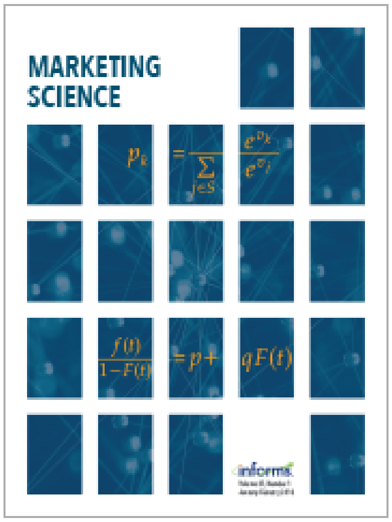


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Testing Models of Strategic Behavior Characterized by Conditional Likelihoods

Thomas Otter, Timothy J. Gilbride, Greg M. Allenby,

MARKETING SCIENCE
CELEBRATING 30 YEARS
Vol. 30, No. 4, July–August 2011, pp. 686–701
ISSN 0732-2399 | EISSN 1526-548X | 11 | 3004 | 0686

informs®
DOI 10.1287/mksc.1110.0644

Testing Models of Strategic Behavior Characterized by Conditional Likelihoods

Thomas Otter
Goethe University, 60054 Frankfurt am Main, Germany, otter@marketing.uni-frankfurt.de

Timothy J. Gilbride
Mendoza College of Business, University of Notre Dame, Notre Dame, Indiana 46556, tgilbrid@nd.edu

Greg M. Allenby
Fisher College of Business, Ohio State University, Columbus, Ohio 43210, allenby_1@osu.edu

Marketing expenditures in the form of pricing, product development, promotion, and channel development are made to maximize profits. A challenge in evaluating the effectiveness of these expenditures is that decisions such as whether to lower prices or run promotions are made based on managers' knowledge of how sensitive consumers are to these marketing activities. Although marketing control variables are explanatory of sales, they are often set in anticipation of a market response, which reflects strategic behavior on the part of a firm. A challenge in developing a model of strategic behavior is that the process by which marketing expenditures are made is often not directly observable. We propose tests for comparing supply-side model formulations in which input variables are strategically determined. In these models, the joint likelihood of demand (y) and supply (x) can be factored into a conditional factor of demand given supply and into a marginal factor of supply. We illustrate our approach using data from a services company that operates in multiple geographic regions.

Key words: structural models; Bayesian analysis; Bayes factors; model selection
History: Received: November 30, 2009; accepted: January 24, 2011; Eric Bradlow and then Preyas Desai served as the editor-in-chief and Michel Wedel served as associate editor for this article. Published online in *Articles in Advance* June 6, 2011.

1. Introduction

A growing literature on the supply-side behavior of firms recognizes the endogeneity of marketing variables. Expenditures in support of product, promotion, price, and channel initiatives are often made by firms to maximize profits by responding to individual wants, attracting new customers, and increasing sales and profits. Marketing mix variables become endogenous whenever they are determined, in part, by parameters that are present in the demand model. As pointed out by many authors, the endogeneity of explanatory variables such as prices can lead to inconsistent parameter estimates in standard models that simply condition on marketing mix variables (Chamberlain 1984, Villas-Boas and Winer 1999, Bronnenberg and Mahajan 2001, Manchanda et al. 2004, Dong et al. 2009).

An elegant way to address this endogeneity problem motivates the observed realizations of explanatory marketing variables through a model reflective of firms' goal-directed behavior. We generically refer to the behavior of firms as the "supply-side" model or simply as the supply model. In these models, a firm's goal-directed behavior is assumed to proceed conditional on knowledge about unobserved demand-generating parameters formally linking the demand and supply models. The econometric endogeneity problem can be resolved by jointly fitting the demand and supply models (e.g., Reiss and Wolak 2007).

Models that include a supply-side story have made significant inroads in marketing. It is hard to embrace the premise that marketing mix variables are set randomly as in an experiment or by a process that is independent of expected outcomes. However, these models rely on very specific assumptions about the knowledge set and optimizing behavior of firms when they set prices, levels of advertising, etc. The resulting models are frequently elegant and parsimonious while providing measures of economic interest. In principle, a direct empirical investigation of managers' decision protocols and knowledge sets is desirable to support or critically question the assumptions made. Without this information, the supply-side model needs to be carefully tested. Testing of the supply-side model is crucial for the research to have empirical content beyond the calibration of parameters for a model assumed to be true a priori.

However, the evaluation of supply-side models is difficult for two reasons: (i) data are limited, and (ii) tests have concentrated on joint fit of the demand and supply data, not on the congruence of these data for explaining common parameters. In this paper, we develop a Bayesian test of the supply-side model, building on the intuition that if the supply-side model is correct, then its inclusion should not only lead to more plausible parameter estimates but also more accurate predictions than using the demand model only. Our test facilitates the statistical comparison of models with and without a supply-side specification and automatically accounts for the fact that the presence of a supply-side model necessarily reduces the maximized fit of the demand-side data. The ultimate test for any model, of course, is whether the effects from causal interventions (such as changes in demand after an exogenous price change) are correctly predicted.[1] However, the more we can say about different supply-side formulations before moving to the ultimate test of implementing a derived policy recommendation, the more we can hope to contribute through research before costly experiments.

The remainder of this paper is organized as follows. Section 2 discusses the scope of our test and gives modeling examples. Section 3 discusses Bayesian tools for model testing. Performance of the methods developed is illustrated in §4 through a simulation study. Section 5 applies the methods to data from a services firm that operates in multiple geographic areas. Concluding remarks are offered in §6.

2. Models of Strategic Behavior with Conditional Likelihoods

Our tests are developed for models where the likelihood of inputs x (i.e., supply) and outputs y (i.e., demand) can be factored into a conditional likelihood of outputs given inputs and into a marginal likelihood of the inputs:

$$\pi(y, x \mid \theta_x, \theta_y, \theta_{x,y}) = \pi(y \mid x, \theta_y, \theta_{x,y}) \times \pi(x \mid \theta_x, \theta_{x,y}), \qquad (1)$$

where θ_y are parameters uniquely found in the demand equation, θ_x are parameters found only in the supply equation, and $\theta_{x,y}$ are shared parameters. An example of θ_y is the scale parameter of the error term in the demand equation, whereas $\theta_{x,y}$ might be a coefficient reflecting price sensitivity. Equation (1) implies that the model for the inputs, or supply-side variables (x), depends on the expectation of demand as predicted by model parameters (θ_x, $\theta_{x,y}$) and not on their realized value (y). This factorization is typically *not* present in models where inputs (x) and outputs (y) are simultaneously determined, as Equation (1) implies a sequencing of variable determination—first x is determined, and then y is realized given x.

Strategic behavior means that the likelihood of supply, $\pi(x \mid \theta_x, \theta_{x,y})$, reflects some goal-directed behavior on the part of management (Chintagunta et al. 2006) such as that encountered in the maximization of profits. In general, the supply variables are marketing variables (4Ps) invested to maximize immediate profits or some form of infrastructure that makes this goal easier in the future.

The factorization in Equation (1) is present in empirical industrial organization models (see Yang et al. 2003), where $\pi(x \mid \theta_x, \theta_{x,y})$ is implied, e.g., by an equilibrium resulting from rational competitive behavior conditional on public and private information that includes knowledge about demand-generating parameters. However, in the following we will confine our discussion to monopolistic situations where marketing control variables are set in anticipation of sales and other outputs without taking competitive reactions into account (e.g., Manchanda et al. 2004). We will revisit these models in the concluding section.

2.1. Monopolist Pricing

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 joint likelihood for the data is a combination of a traditional demand model:

$$\ln y_t = \beta_0 + \beta_1 \ln p_t + \varepsilon_t; \quad \varepsilon_t \sim \mathrm{N}(0, \sigma_\varepsilon^2) \qquad (2)$$

and a factor for the endogenous price variable. Profit-maximizing pricing for a monopolist can be shown to be

$$p_t = mc\left(\frac{\beta_1}{1+\beta_1}\right)e^{v_t}; \quad v_t \sim \mathrm{N}(0, \sigma_v^2), \qquad (3)$$

where *mc* denotes the marginal cost of the brand, and a supply-side error term has been added to account for a temporal variation of observed prices from the optimal price. In this example, $\theta_y' = (\beta_0, \sigma_\varepsilon^2)$, $\theta_x = \sigma_v^2$, and $\theta_{x,y} = \beta_1$.

The supply-side error term in Equation (3) has a structural interpretation as a (perceived) cost shifter. Observed prices are set as if marginal costs varied over time, suggesting that the supply-side error realizations are known to the decision maker. However, the supply-side errors are not part of the data and thus are unobservable to the analyst. Taking logs of Equation (3) yields

$$\ln p_t = \ln mc + \ln\left(\frac{\beta_1}{1+\beta_1}\right) + v_t; \quad v_t \sim \mathrm{N}(0, \sigma_v^2). \qquad (4)$$

[1] See Misra and Nair (2009) for an impressive validation of a structural model through the actual implementation of a derived policy.

The supply-side equation (4) can be used to help locate the value of β_1 if the marginal cost is known. It also implies, through the second-order conditions of the optimization problem, ordinal constraints on the model parameters—i.e., $\beta_1 < -1$ for a positive, finite price at a profit maximum. Optimal pricing behavior for the monopolist only exists when own-price effects are elastic. The supply-side equation (4), derived from first-order conditions, influences estimation in line with assuming a profit maximum at positive, finite prices, even when marginal costs are not known. This is because the error term is structurally confined in that it cannot be used to explain managers' lack of knowledge about the response coefficients, such as mistaking inelastic demand for elastic demand. From Equation (3), we can see that the error term v_t can only expand or contract the ratio $B_1/(1+B_1)$. The supply-side equation can therefore impart a great deal of information for decision making, even when costs are not observed. It is therefore important to have methods for developing and assessing the validity of supply-side models.

Equations (3) and (4) are not the only explanation of how prices could be set in this stylized example. An alternative and more common formulation involves the presence of an omitted variable, η_t, that creates correlation between demand and prices:

$$\begin{gathered}\ln y_t = \beta_0 + \beta_1 \ln p_t + (\eta_t + \varepsilon_t^*), \quad \varepsilon_t = \eta_t + \varepsilon_t^*,\\ p_t = mc\left(\frac{\beta_1}{1+\beta_1}\right) e(\eta_t + v_t^*), \quad v_t = \eta_t + v_t^*,\\ \begin{pmatrix}\varepsilon_t\\ v_t\end{pmatrix} \sim \mathrm{N}(0, \Sigma).\end{gathered} \tag{5}$$

The omitted variable η_t increases demand and marginal costs exponentially and is assumed known in advance to the agent setting prices. In this model, price variation is due to the omitted variable as opposed to just variation in marginal costs as in (3). For example, demand for hotel rooms increases exponentially over the weekend of a popular football game and so do marginal costs because of overtime for staff. In this situation, the hotel can pass on these higher costs through higher room rates in anticipation of a peak in demand.

As shown by Yang et al. (2003), by conditioning on η_t, we can write the likelihood as in Equation (1), and the Bayesian method of data augmentation can then be used to integrate over the distribution of η_t. Thus, certain models traditionally viewed as simultaneous can often be recast with a conditional likelihood function considered in this paper. The key difference is whether prices are set in anticipation of a market outcome versus being driven by realized demand. The former implies a recursive system between the supply-side variables and demand—e.g., η_t reflects information used to set prices—and once prices are set, demand is then realized.

A third explanation for variation in prices over time could be due to anticipated variation in price elasticity through time:

$$\begin{gathered}\ln y_t = \beta_0 + \beta_{1,t} \ln p_t + \varepsilon_t,\\ p_t = mc\left(\frac{\beta_{1,t}}{1+\beta_{1,t}}\right) e(v_t),\\ \begin{pmatrix}\varepsilon_t\\ v_t\end{pmatrix} \sim \mathrm{N}(0, \mathbf{I}[\sigma, \sigma]),\end{gathered} \tag{6}$$

where $\mathbf{I}$ represents the 2×2 identity matrix, and we note the time (t) subscript for the price coefficient. From Equations (3)–(6), it is evident that multiple goal-directed explanations for the data exist, each with possibly different implications about optimal price-setting behavior. Price variation as a result of random variation in costs leads to different prices but the same percentage markup, whereas price variation as a result of variation in the price elasticity itself implies variation in the optimal markup. Our goal is to develop ways to understand which explanation of the data is correct, including the possibility that none of the goal-directed formulations under investigation applies.

2.2. Cobb-Douglas Production and Optimal Allocation

For a supply-side model to be goal directed, it must involve aspects of the demand model as reflected though the shared parameters $\theta_{x,y}$ and the first-order conditions associated with some criterion function (e.g., profits) that the firm wishes to maximize. As a second example, we develop a constrained allocation problem in the context of a Cobb-Douglas production function, allowing us to consider multiple interacting input variables.

Many authors have used the Cobb-Douglas function in marketing—Misra et al. (2005) posits a hierarchical model structure to account for district and regional effects in decomposing salesforce technical effects, Horsky and Nelson (1996) propose a method for benchmarking sales response, Carroll et al. (1979) investigate preferences for leisure time, Morey and McCann (1983) study optimal lead generation from advertising, and Mantrala et al. (1992) 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 on the same model parameters, $\theta_{x,y}$. Thus, these previous methods that condition on the inputs are prone to misspecification. The models illustrated next resolve the simultaneity by incorporating the first-order conditions for optimally allocating marketing expenditures. Outside of marketing, researchers have used variants of the Cobb-Douglas production function together with first-order conditions to assess technical and allocative efficiency (Zellner et al. 1966).

We begin by considering general first-order conditions. Optimal allocation 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 (x^*) can be determined by forming the auxiliary production function (L) that includes the budget constraint and Lagrangian multiplier (λ), by taking derivatives, and by solving the system of equations that arises 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:

$$\frac{\partial y/\partial x_j}{p_j} = \frac{\partial y/\partial x_k}{p_k} \quad \text{for all } j \text{ and } k, \tag{7}$$

where the subscripts j and k refer to different input variables, and p_j denotes the price of input j. Equation (7) can be used to specify a system of equations that reflect goal-directed behavior, and it pertains to situations where the objective function is proportional to the output measure. Consider a standard Cobb-Douglas production function and associated first-order conditions:

$$\begin{aligned} y &= \beta_0 x_1^{\beta_1} x_2^{\beta_2}, \\ \lambda &= \frac{\partial y/\partial x_1}{p_1} = \frac{\beta_1}{p_1}\beta_0 x_1^{\beta_1 - 1} x_2^{\beta_2}, \\ \lambda &= \frac{\partial y/\partial x_2}{p_2} = \frac{\beta_2}{p_2}\beta_0 x_1^{\beta_1} x_2^{\beta_2 - 1}. \end{aligned} \tag{8}$$

Taking logs and rearranging terms, we have

$$\begin{aligned} &\ln(y) = \ln\beta_0 + \beta_1\ln(x_1) + \beta_2\ln(x_2), \\ &\ln(x) = \begin{bmatrix} \ln(x_1) \\ \ln(x_2) \end{bmatrix} = \begin{bmatrix} \beta_1 - 1 & \beta_2 \\ \beta_1 & \beta_2 - 1 \end{bmatrix}^{-1} \\ &\qquad \cdot \begin{bmatrix} \ln\lambda - \ln\beta_0 - \ln\beta_1 + \ln p_1 \\ \ln\lambda - \ln\beta_0 - \ln\beta_2 + \ln p_2 \end{bmatrix}. \end{aligned} \tag{9}$$

We note that Equation (9) is not the only model of strategic behavior that could be considered. Our goal is to test alternative models, including the possibility of the simple conditional demand model.

2.3. Structural Error Terms

Observed deviations from optimal allocations are accommodated by associating error with the optimal allocation rule in (7). For example, errors (ζ) could be structurally associated with input prices, i.e., $p_j^* = p_j e^{\zeta_j}$, assuming that Equation (7) holds exactly for p_j replaced by p_j^*. In this case, the unobservable (ζ) is an argument in the optimization problem, suggesting that (ζ) is observed by the decision maker and is open to a structural interpretation.

Updating Equation (9) by adding additive observation[2] error (ε_i) to the demand equation, and assuming that inputs are allocated based on $p_i^* = pe^{\zeta_i}$ rather than p, leads to estimation equations:

$$\begin{aligned} &y_i = \beta_0 x_{1,i}^{\beta_1} x_{2,i}^{\beta_2} + \varepsilon_i, \\ &\ln(x_i) = \begin{bmatrix} \ln(x_{1i}) \\ \ln(x_{2i}) \end{bmatrix} = \begin{bmatrix} \beta_1 - 1 & \beta_2 \\ \beta_1 & \beta_2 - 1 \end{bmatrix}^{-1} \\ &\qquad \cdot \begin{bmatrix} \ln\lambda - \ln\beta_0 - \ln\beta_1 + \ln p_1 + \zeta_{1i} \\ \ln\lambda - \ln\beta_0 - \ln\beta_2 + \ln p_2 + \zeta_{2i} \end{bmatrix}. \end{aligned} \tag{10}$$

This system is just identified for known λ and prices, and repeated observations 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. Variation in the input variables is entirely due to factors known to the decision maker but not observed by the analyst (i.e., allocation errors). By rearranging the equation that generates the inputs and applying the change of variable theorem, we see that the solution space for the parameters of the production function is bounded away from the region of constant marginal returns for any distribution of (ζ):

$$\begin{aligned} \zeta_i &= \begin{bmatrix} \zeta_{1i} \\ \zeta_{2i} \end{bmatrix} = \begin{bmatrix} \beta_1 - 1 & \beta_2 \\ \beta_1 & \beta_2 - 1 \end{bmatrix} \begin{bmatrix} \ln(x_{1i}) \\ \ln(x_{2i}) \end{bmatrix} \\ &\qquad - \begin{bmatrix} \ln\lambda - \ln\beta_0 - \ln\beta_1 + \ln p_1 \\ \ln\lambda - \ln\beta_0 - \ln\beta_2 + \ln p_2 \end{bmatrix}, \\ &|J_{\zeta\to\ln x,\, i}| = |(\beta_1 - 1)(\beta_2 - 1) - \beta_1\beta_2|, \\ &\pi(\{\ln x_i\} \mid \cdot) = \pi(\{\zeta_i\} \mid g)|J_{\zeta\to\ln x_i}|. \end{aligned} \tag{11}$$

The Jacobian term evaluates to zero for constant marginal returns. This rules out continuous moves from the solution space with diminishing marginal returns to the solution space with increasing marginal returns. Also, different from the simple ordinal constraint $0 < \beta_1 + \beta_2 < 1.0$, the density of positive allocations decreases continuously in the direction of either boundary.

It is important to note that a flat prior for the allocation error does not automatically translate into a flat

[2] We use additive instead of the more common multiplicative error structure. The multiplicative error structure is often chosen to facilitate estimation. However, in most marketing applications, zero inputs (e.g., no advertising in the current period) do not necessarily translate into the zero output implied by multiplicative errors. Additive errors capture this fact.

posterior from the supply-side model. For example, the likelihood for the inputs implied by the structural errors ζ in Equation (11) becomes informative as the response coefficients approach constant returns to scale; i.e., $\beta_1+\beta_2=1$. For a flat distribution $\pi(\zeta_i)$, the likelihood of the log inputs is proportional to the Jacobian term with a minimum ridge corresponding to the set of parameters that imply constant returns to scale:

$$\pi(\{\ln x_i\}\mid\cdot)\propto|(\beta_1-1)(\beta_2-1)-\beta_1\beta_2|. \tag{12}$$

The structural interpretation of an unobservable is tightly linked to what cannot be explained by any possible realization of that unobservable. This was illustrated with the monopolist pricing example, where no realization of v_t in Equation (3) could lead to managers mistaking inelastic demand for elastic demand. Replacing the input prices p by p^* in Equation (7) retains the assumption that an interior allocation optimum exists and thus puts zero prior probability on managers mistaking constant or increasing marginal returns for decreasing marginal returns.

This is an example of how some measure of "rationality" is assumed in the model. A structural supply-side model helps locate parameters within the region that is consistent with interior optimal solutions. We develop a test to determine to what extent the assumed measure of rationality is empirically supported by the data that complement essentially a priori arguments based on face validity. However, the question of whether the unobservables (ζ) reflect allocation errors or material private information held by the managers has to be judged on grounds other than observed inputs (x), outputs (y), and empirical fit.

2.4. Descriptive Supply Models

An alternative approach is to formulate a descriptive supply-side specification that can impart minimal influence on common parameters $\theta'_{x,y}=(\beta_0,\beta_1,\beta_2)$ if the data indicate. Such an approach is proposed by Manchanda et al. (2004), who introduce additional parameters in the supply-side model that, when zero, nullify the effect of the supply side on $\theta_{x,y}$. It turns out, however, that this result (i.e., a descriptive supply model being directionally consistent with output increases) is rather unique and is driven by the chosen response function, which has only one input and nondecreasing marginal returns. When multiple inputs are present and interact with each other, the optimal allocation of a single input does not necessarily increase with an increase in output.

Consider three units producing according to different Cobb-Douglas specifications:

$$\begin{aligned} y_1&=x_{11}^{0.6}x_{21}^{0.1},\\ y_2&=x_{12}^{0.1}x_{22}^{0.6},\\ y_3&=x_{13}^{0.45}x_{23}^{0.45}.\end{aligned} \tag{13}$$

Table 1 **Optimal Allocation to Three Heterogeneous Cobb-Douglas Production Units**

	$\lambda=0.4$		$\lambda=0.44$		$\lambda=0.5$	
Unit	Input 1	Input 2	Input 1	Input 2	Input 1	Input 2
1	2.13	0.35	1.55	0.26	1.01	0.17
2	0.35	2.13	0.26	1.55	0.17	1.01
3	3.25	3.25	1.25	1.25	0.35	0.35

Setting input prices to 1, we obtain the optimal allocations in Table 1 for less and more attractive outside investment options, represented by smaller and larger Lagrangian multipliers (λ).

With the least attractive outside option ($\lambda=0.4$) investigated here, unit 3 in Table 1 receives the highest allocations along both inputs even though it does not have the (individually) highest coefficients. When the attractiveness of the outside option is increased ($\lambda=0.44$), the optimal allocation suggests a different ordering. Now, unit 1 receives the highest input 1 and unit 2 the highest input 2. The sum of the two inputs is still maximal for unit 3. When the attractiveness of the outside option increases further ($\lambda=0.5$), the sum of the two inputs is now the smallest for unit 3. This simple example illustrates the difficulty in extending the notion of a descriptive supply-side model of strategic behavior where the inputs interact.

We develop an empirical test of the supply-side formulation for this and similar situations where a directionally consistent descriptive supply-side formulation is not available and where a (necessarily) directionally consistent supply-side model derived from first-order conditions cannot revert to a neutral state of zero influence on the demand side.

3. Bayesian Tests of Strategic Behavior

In this section, we discuss Bayesian tests of strategic behavior for likelihoods of inputs (x) and outputs (y) that can be factored as in Equation (1). A simple approach to drawing inference about the supply-side model is to compare predictive densities from posteriors informed by different subsets of the data. One comparison would be the predictive fit of holdout demand data (y^h) using the supply-side variables (x^h, $\{x\}$) to help inform the posterior distribution, i.e., $\int\pi(y^h\mid x^h,\theta)\pi(\theta\mid y)\,d\theta$ versus

$$\int\pi(y^h\mid x^h,\theta)\pi(\theta\mid y,x)\,d\theta \quad\text{and}$$

$$\int\pi(y^h\mid x^h,\theta)\pi(\theta\mid y,x,x^h)\,d\theta$$

(e.g., Manchanda et al. 2004). The model for the inputs is supported if

$$\begin{aligned}&\int\pi(y^h\mid x^h,\theta)\pi(\theta\mid y)\,d\theta\\ &\quad<\int\pi(y^h\mid x^h,\theta)\pi(\theta\mid y,x)\,d\theta\\ &\quad<\int\pi(y^h\mid x^h,\theta)\pi(\theta\mid y,x,x^h)\,d\theta.\end{aligned} \tag{14}$$

That is, predictions based on a demand model are worse than predictions that successively use more information about θ gleaned from the supply-side model. Although this approach can be used to compare different prior specifications for the supply side including the special case without a supply side, it requires sufficient data so that a reasonably sized holdout data set can be created. For example, Dong et al. (2009) estimate their model with 11 consecutive quarters of individual-level data, leaving only the 12th for the holdout sample. We view the possibility of a direct comparison to an agnostic conditional model as essential because the joint fit of the demand and the allocations may be masking the event that the allocation model is negatively affecting the fit of the response data. We now propose a way to achieve this comparison in small data sets that do not allow for extensive holdout samples.

We introduce a new method of testing by first specifying the prior distribution of model parameters in (1) as having independent factors that facilitate the comparison of the joint model to the conditional demand model:

$$\pi(\theta_y, \theta_x, \theta_{x,y}) = \pi(\theta_y)\pi(\theta_x)\pi(\theta_{x,y}). \tag{15}$$

The joint marginal likelihood of inputs and outputs in (1) is then

$$\begin{aligned}\int \pi(y \mid x, \theta_y, \theta_{x,y})\pi(x \mid \theta_x, \theta_{x,y})\pi(\theta_y)\pi(\theta_x) \\ \cdot\pi(\theta_{x,y})\, d\{\theta_y, \theta_x, \theta_{x,y}\} = \pi(y, x \mid M_f) \\ = \pi(y \mid x, M_f)\pi(x \mid M_f),\end{aligned} \tag{16}$$

and the conditional marginal likelihood is given by

$$\begin{aligned}&\int \pi(y \mid x, \theta_y, \theta_{x,y})\pi(\theta_y)\pi(\theta_{x,y})\, d\{\theta_y, \theta_{x,y}\} \\ &\quad = \pi(y \mid x, M_c).\end{aligned} \tag{17}$$

Our goal is to assess the reasonableness of the supply-side specification by comparing an estimate of $\pi(y \mid x, M_f)$ from (16) to $\pi(y \mid x, M_c)$ from (17). Estimates of $\pi(y \mid x, M_f)$ can also be used to assess the relative merits of different supply-side formulations, which gives rise to different joint models M_f. If the supply model is misspecified, then the conditional marginal density of $\pi(y \mid x,\ M_f)$ will be smaller than the corresponding estimate of $\pi(y \mid x, M_c)$. Also, if a particular supply-side formulation is better supported by the demand data, the corresponding $\pi(y \mid x, M_f)$ will be larger than that of another full model derived with a different supply-side story.

We derive two estimators for $\pi(y \mid x, M_f)$. The first provides a direct estimate of $\pi(y \mid x, M_f)$ similar to the Newton and Raftery (1994) estimator. The second is an indirect estimator based on the identity $\pi(y \mid x, M_f) = \pi(y, x \mid M_f)/\pi(x \mid M_f)$, the building blocks of which are estimated following Gelfand and Dey (1994). The advantage of the direct Newton–Raftery (NR) estimator is that it can be very easily computed and understood. A disadvantage is that it can suffer from the same simulation bias as discussed by Lenk (2009) in the context of estimating $\pi(y \mid x, M_c)$. The indirect Gelfand and Dey (GD) estimator is computed as the ratio of two normalizing constants. The advantages of the indirect estimator are that it is more accurate and it is always pointed to the correct model in our simulations (see §4). A disadvantage of the indirect estimator is that it is harder to compute.

The direct NR-style estimator derived in Appendix A is defined as

$$\begin{aligned}\pi(y \mid x, M_f) = \Bigg\{ \int \frac{1}{\pi(y \mid x, \theta_y, \theta_{x,y})} \\ \cdot \frac{\pi(y \mid x, \theta_y, \theta_{x,y})\pi(x \mid \theta_x, \theta_{x,y})\pi(\theta_x)\pi(\theta_y)\pi(\theta_{x,y})}{\pi(y \mid x)\pi(x)} d\{\theta_x, \theta_y, \theta_{x,y}\}\Bigg\}^{-1}.\end{aligned} \tag{18}$$

This results in the following approximation based on an Markov chain Monte Carlo (MCMC) sample from the (marginal) posterior from the full model:

$$\pi(y \mid x, M_f) \approx \left\{ \frac{1}{M} \sum_{m=1}^{M} \pi(y \mid x, \theta_y^{(m)}, \theta_{x,y}^{(m)})^{-1} \right\}^{-1}; \tag{19}$$

that is, the estimator simply averages the inverse of the conditional likelihood of the demand data evaluated at parameter draws from the full model, which includes the supply side.

A practical caveat to this particular estimator of $\pi(y \mid x, M_f)$, not to the underlying theory, is Lenk's (2009) demonstration of simulation bias in the NR approximation because of the practically limited support of MCMC samples from (concentrated) posterior distributions relative to their theoretical support and that of the prior. The proposed estimator will suffer from similar practical problems as documented by Lenk (2009) if the prior on parameters only informed by the demand model, i.e., $\pi(\theta_y)$ in our generic notation, is much more diffuse than its posterior counterpart $\pi(\theta_y \mid y, x)$ or if the information contributed to the posterior by the demand-side model outweighs the information contributed by the supply-side model disproportionally.

We therefore construct an alternative, indirect estimator of the marginal conditional density $\pi(y \mid x, M_f)$ using the identity $\pi(y \mid x, M_f) = \pi(y, x \mid M_f)/\pi(x \mid M_f)$. Reliable estimates of $\pi(y, x \mid M_f)$ and $\pi(x \mid M_f)$ are obtained following Gelfand and Dey (1994). The GD method also relies on the posterior represented in the form of MCMC draws as an integrating density. In addition, it uses the posterior to calibrate suitably thin-tailed normalized importance densities

$q(\theta)$ to be used in the generic expression (see also Rossi et al. 2005, Chapter 6):

$$\pi(data \mid M_i) = \left\{ \int \frac{q(\theta)}{\pi(data \mid \theta, M_i)\pi(\theta \mid M_i)} \pi(\theta \mid data, M_i)\, d\theta \right\}^{-1}. \quad (20)$$

The calibration of the normalized importance densities $q(\theta)$ from the posterior, incorporating possible constraints from the supply side, avoids the practical and theoretical problems of the NR-style estimator. Details for estimating (20) are in Appendix A.

4. Simulation Study

We demonstrate the potential of our proposed test statistic using two simulation studies. The first study involves simulated data from the monopolist pricing example discussed. The second study examines the performance of the test statistic using the Cobb-Douglas function in Equation (10) and anticipates analysis performed in our empirical application.

The first simulation study examines the sensitivity of the test statistics to alternative supply-side specifications. Priors are specified as being relatively diffuse but proper across the models, with the prior on coefficients in Equation (2) set to $\mathrm{N}(0, 100)$ and the prior on variance terms set to Inverted Gamma$(3, 0.1 \times [3-1])$. We investigate four different data-generating mechanisms:

i. *Demand only*. A standard regression model as in Equation (2) with prices generated from Gamma$(30, 15)$ and $\sigma^2_{\varepsilon*} = \sigma^2_{\varepsilon} = 0.1$.

ii. *Demand + Supply, random costs*. A joint model with price variation as a result of observed cost variation distributed Gamma$(30, 15)$, unobserved cost variation in Equation (4) with $\sigma^2_{v*} = \sigma^2_{v} = 0.1$, and the same demand error variance as in (i) above.

iii. *Demand + Supply, common shocks*. A joint model with price variation as a result of observed cost variation as above, plus common shocks as in (5) with a variance of $\sigma^2_{\eta} = 0.50$, implied demand and supply error variances $\sigma^2_{\varepsilon} = \sigma^2_{\varepsilon*} + \sigma^2_{\eta} = 0.6$, $\sigma^2_{v} = \sigma^2_{v*} + \sigma^2_{\eta} = 0.6$, and an implied correlation of 0.83 between demand- and supply-side errors.

iv. *Demand + Supply, time-varying elasticity*. A joint model with price variation as a result of observed cost variation as in (ii) above, temporal random effects as in (6) with $\beta_{1,t} \sim \mathrm{N}(-2, 0.15)$, and demand and supply error variances $\sigma^2_{\varepsilon*} = \sigma^2_{\varepsilon} = 0.1$, $\sigma^2_{v*} = \sigma^2_{v} = 0.1$.

We simulate a year's worth of weekly data from each data-generating mechanism and fit four models to each data set: (i) a demand-only model, (ii) a demand and supply model with random cost shocks, (iii) a demand and supply model with common shocks on top of random cost components, and (iv) a demand and supply model with time-varying elasticities indentified through the supply side. We collect parameter estimates from all models fitted to all of the four data-generating mechanisms in Table 2. The corresponding test statistics are presented in Table 3.

When the data-generating mechanism matches the model, on the diagonal of Table 2, parameters are correctly recovered. When the data are generated from a conditional model where the prices do not contain any information about unobservables in the demand model, adding any of the three candidate supply-side models results in severe biases. Note that the data-generating price elasticity is -0.7 and in the inelastic range, but adding any of the three supply-side models forces the elasticity into the elastic range. Without access to the formal test we propose, we believe that many researchers would be tempted to accept one of the joint models based on presumed plausibility of elastic demand.

When the data are generated from a joint model, fitting the conditional model that only considers the demand side will result in elasticities that are (absolutely) too small. The exception is the joint model, where only random cost shocks cause temporal price variation. With this data-generating mechanism, consistent inferences from a demand-only model are possible because parameters driving price variation are orthogonal to parameters in the demand model.

Table 3 shows that the proposed test statistic $\ln \pi(y \mid x, M)$, estimated from MCMC output using the indirect GD estimator, identifies the correct model in every case (along the diagonal of Table 3). It correctly rejects all supply-side models when the data are generated from a simple conditional model, debunking the increased face validity from adding any of the supply-side stories. It correctly rejects the demand-only model when the data were generated with a supply side and identifies the data-generating formulation. When consistent inferences are possible from the demand-only model such as in the case of random cost shocks, the test statistic correctly picks up the increased efficiency from adding the correctly specified supply-side.

Comparing estimates from the direct (NR) and the indirect estimator (GD), it seems that the direct estimator may be too optimistic, which is in line with the results in Lenk (2009). More importantly, the direct (NR) estimator of $\ln \pi(y \mid x, M)$ incorrectly points to the model with time-varying elasticity, which is much larger in dimensionality when the data are generated with random costs.

Our second simulation study investigates whether the proposed test statistic can detect the (lack of) contribution of a formal allocation model in the context of Cobb-Douglas production functions. We investigate three different data-generating mechanisms. In

Table 2 Parameter Estimates: Monopolist Pricing Example, Posterior Mean

Parameters	True values	Demand only	Demand + Supply, random costs	Demand + Supply, common shocks	Demand + Supply, time-varying elasticity
		Data generation: Demand only			
β_n	1.00	0.95 (0.119)[a]	2.26 (0.192)	2.72 (0.105)	2.37 (0.193)
β_1	−0.70	−0.58 (0.145)	−2.30 (0.231)	−2.29 (0.168)	−2.50 (0.243)
Var[β_1]	0.00	NA	NA	NA	0.11 (0.064)
Var[ε^*]	0.10	0.10 (0.019)	0.37 (0.106)	0.15 (0.052)	0.36 (0.110)
Var[v^*]	NA	NA	0.41 (0.116)	0.06 (0.028)	0.35 (0.096)
Var$\|\eta\|$	NA	NA	NA	0.41 (0.099)	NA
Corr[ε, v]	NA	NA	NA	0.79 (0.058)	NA
		Data generation: Demand + Supply, random costs			
β_n	1.00	0.91 (0.193)	0.91 (0.127)	1.22 (0.142)	1.00 (0.111)
β_1	−2.00	−2.01 (0.137)	−2.01 (0.084)	−2.30 (0.120)	−2.09 (0.087)
Var[β_1]	NA	NA	NA	NA	0.06 (0.016)
Var[ε^*]	0.10	0.14 (0.027)	0.17 (0.026)	0.10 (0.030)	0.06 (0.023)
Var[v^*]	0.10	NA	0.12 (0.023)	0.08 (0.026)	0.12 (0.024)
Var$\|\eta\|$	NA	NA	NA	0.11 (0.030)	NA
Corr[ε, v]	NA	NA	NA	0.57 (0.078)	NA
		Data generation: Demand + Supply, common shocks			
β_n	1.00	−0.14 (0.117)	0.38 (0.124)	1.05 (0.095)	0.49 (0.111)
β_1	−2.00	−1.11 (0.084)	−1.55 (0.076)	−2.00 (0.085)	−1.71 (0.091)
Var[β_1]	NA	NA	NA	NA	0.08 (0.029)
Var[ε^*]	0.10	0.22 (0.043)	0.33 (0.074)	0.14 (0.046)	0.25 (0.064)
Var[v^*]	0.10	NA	0.78 (0.174)	0.08 (0.042)	0.56 (0.129)
Var$\|\eta\|$	0.50	NA	NA	0.53 (0.113)	NA
Corr[ε, v]	0.83	NA	NA	0.83 (0.041)	NA
		Data generation: Demand + Supply, time-varying elasticity			
β_n	1.00	0.19 (0.197)	0.80 (0.128)	1.16 (0.130)	1.02 (0.08)
β_1	−2.00	−1.33 (0.136)	−1.78 (0.079)	−2.06 (0.117)	−2.01 (0.077)
Var[β_1]	0.15	NA	NA	NA	0.11 (0.027)
Var[ε^*]	0.10	0.21 (0.041)	0.25 (0.051)	0.12 (0.035)	0.05 (0.017)
Var[v^*]	0.10	NA	0.26 (0.051)	0.07 (0.028)	0.13 (0.028)
Var$\|\eta\|$	NA	NA	NA	0.22 (0.054)	NA
Corr[ε, v]	NA	NA	NA	0.70 (0.066)	NA

[a]Posterior standard deviation.

Table 3 Test Statistics: Monopolist Pricing Example

		Demand only	Demand + Supply, random costs	Demand + Supply, common shocks	Demand + Supply, time-varying elasticity
		Data generation: Demand only			
$\ln \pi(y \mid x, M)$	GD	−23.37	−93.77	−71.63	−94.32
	NR	−15.61	−67.44	−43.06	−60.67
		Data generation: Demand + Supply, random costs			
$\ln \pi(y \mid x, M)$	GD	−34.06	−28.84	−32.19	−29.39
	NR	−25.38	−24.30	−25.44	−18.81
		Data generation: Demand + Supply, common shocks			
$\ln \pi(y \mid x, M)$	GD	−48.14	−66.11	−41.33	−63.59
	NR	−37.48	−53.27	−34.01	−49.36
		Data generation: Demand + Supply, time-varying elasticity			
$\ln \pi(y \mid x, M)$	GD	−74.06	−77.13	−70.65	−56.99
	NR	−62.90	−68.63	−61.98	−38.42

the first data-generating mechanism, the observed allocation does not contain any additional information about unobservables in the demand model (the allocation is exogenous). The second data-generating mechanism allocates endogenously following Equation (10). However, the *variance* in the allocated inputs is exogenously driven by independent supply-side errors only, i.e., there are no missing input factors, so that the true model has a fixed, non-heterogeneous intercept. Finally, the third data-generating mechanism has inputs optimally allocated according to Equation (10), with near-zero allocation error—i.e., $V(\zeta_i) = 1.0 \times 10^{-6} I$ and differences between production units through heterogeneity in the multiplicative constant unobserved by the analyst. We simulate 21 observations from each of the three data-generating mechanisms, the same number of observations as in our empirical application reported in this paper.

Table 4 Parameter Estimates: Cobb-Douglas Allocation Example

Parameters	True values	Demand only	Demand + Supply fixed effects	Demand + Supply random effects
		Simulated data set 1—Random allocation		
$\ln\beta_1$	−1.50	−1.55 (0.031)	−1.54 (0.032)	−1.58 (0.063)
$\ln\beta_2$	−1.50	−1.57 (0.044)	−1.57 (0.044)	−1.64 (0.100)
$E[\beta_0]$	5.00	5.38 (0.208)	5.38 (0.209)	5.65 (0.352)
$\mathrm{Var}[\beta_0]$	0.00	NA	NA	0.33 (0.133)
$\ln\lambda_1$	NA	NA	−1.28 (0.347)	−1.32 (0.353)
$\ln\lambda_2$	NA	NA	−0.98 (0.308)	−1.05 (0.323)
$\mathrm{Var}[\zeta_1]$	NA	NA	2.54 (0.839)	2.53 (0.823)
$\mathrm{Cov}[\zeta_1, \zeta_2]$	NA	NA	−0.78 (0.536)	−0.78 (0.546)
$\mathrm{Var}[\zeta_2]$	NA	NA	1.96 (0.636)	2.00 (0.648)
σ_ε^2	1.00	0.62 (0.196)	0.62 (0.197)	0.59 (0.283)
		Simulated data set 2—Strategic allocation		
$\ln\beta_1$	−1.50	−1.48 (0.119)	−1.46 (0.120)	−1.46 (0.119)
$\ln\beta_2$	−1.50	−1.39 (0.097)	−1.38 (0.089)	−1.39 (0.091)
$E[\beta_0]$	5.00	4.61 (0.225)	4.58 (0.224)	4.60 (0.225)
$\mathrm{Var}[\beta_0]$	NA	NA	NA	0.66 (0.405)
$\ln\lambda_1$	−0.35	NA	−0.32 (0.219)	−0.32 (0.218)
$\ln\lambda_2$	0.05	NA	0.24 (0.227)	0.23 (0.228)
$\mathrm{Var}[\zeta_1]$	1.00	NA	0.76 (0.254)	0.76 (0.260)
$\mathrm{Cov}[\zeta_1, \zeta_2]$	0.00	NA	−0.31 (0.209)	−0.31 (0.215)
$\mathrm{Var}[\zeta_2]$	1.00	NA	0.96 (0.317)	0.96 (0.327)
σ_ε^2	1.00	0.77 (0.243)	0.76 (0.240)	0.76 (0.238)
		Simulated data set 3—Efficient allocation		
$\ln\beta_1$	−1.50	0.09 (0.082)	−0.54 (0.460)	−1.50 (0.001)
$\ln\beta_2$	−1.50	−8.97 (1.37)	−1.58 (0.460)	−1.50 (0.002)
$E[\beta_0]$	4.80	2.92 (0.027)	3.89 (0.395)	4.75 (0.011)
$\mathrm{Var}[\beta_0]$	2.00	NA	NA	1.86 (0.553)
$\ln\lambda_1$	−0.35	NA	0.65 (0.460)	−0.34 (0.009)
$\ln\lambda_2$	0.05	NA	0.02 (0.88)	0.06 (0.010)
$\mathrm{Var}[\zeta_1]$	1.0×10^{-6}	NA	0.003 (0.003)	1.03×10^{-6}
$\mathrm{Cov}[\zeta_1, \zeta_2]$	0	NA	0.003 (0.003)	-0.002×10^{-8}
$\mathrm{Var}[\zeta_2]$	1.0×10^{-6}	NA	0.003 (0.003)	1.03×10^{-6}
σ_ε^2	1.00	1.04 (0.323)	1.20 (0.371)	1.02 (0.297)

We fit three models to each data set: (i) a demand-model only, (ii) a demand and supply model assuming homogeneous production functions, and (iii) a demand and supply model with random multiplicative intercepts indentified through the supply side, i.e., the allocation model. All subjective priors are standard proper conjugate priors chosen to be minimally informative unless noted otherwise. The model comparison is challenging because of the small number of observations. Table 4 collects parameter estimates, and Table 5 presents test statistics.

Parameters are correctly recovered when the data-generating mechanism matches the estimated model. Parameter estimates are severely biased when "Efficient allocation" based on the knowledge of random multiplicative intercepts is ignored.[3]

When variance in the allocation is driven by exogenous allocation errors ("Strategic allocation"), the demand-only model is consistent, which is confirmed by our results.

Table 5 Test Statistics: Cobb-Douglas Allocation Example

		Demand only	Demand + Supply fixed effects	Demand + Supply random effects
			Simulated data set 1—Random allocation	
$\ln\pi(y \mid x, M)$	GD	−147.6	−148.18	−187.91
	NR	−26.41	−26.47	−35.06
			Simulated data set 2—Strategic allocation	
$\ln\pi(y \mid x, M)$	GD	−151.95	−140.47	−150.15
	NR	−29.25	−29.11	−29.2
			Simulated data set 3—Efficient allocation	
$\ln\pi(y \mid x, M)$	GD	−149.39	−147.8	−116.18
	NR	−32.4	−34.03	−31.07

Note. With a standard diffuse prior $\pi(V_\zeta) = \mathrm{IW}(2, 0.5 \times 10^{-3}I)$, we obtain the following fits when the allocation is "efficient" and a joint model with random effects is estimated: $\ln\pi(y \mid x, M)$: (GD), −134.5; and NR, −32.28.

[3] This level of parameter recovery reported for the efficient allocation with random effects is only achieved with an informative prior, $\pi(V_\zeta) = \mathrm{IW}(51, 0.495 \times 10^{-4}I)$, in line with very small allocation errors.

Again, the proposed test statistic $\ln \pi(y \mid x, M)$ reported in Table 5 identifies the correct model in every case (along the diagonal of Table 5). When the allocation is not driven by unobserved differences in responsiveness across production units ("Random allocation" and "Strategic allocation"), the proposed test clearly rejects the models that identify a random intercept through the supply-side formulation. When the allocation is endogenous ("Strategic allocation" and "Efficient allocation"), we find that the addition of the *supply* side actually increases the conditional marginal likelihood of the *demand* data. When the allocation is endogenous but its variance is exogenous ("Strategic allocation"), the proposed test statistic picks up the increased efficiency from adding the supply model.[4]

Finally, when the allocation is random and thus exogenous, the test statistic barely differentiates between the joint model with fixed effects and the demand-only model. This is not a problem with the test statistic but a feature of the data-generating mechanism. With the data-generating parameters in line with decreasing marginal returns and unobserved input prices, an allocation model driven by exogenous allocation errors cannot fit the demand data any worse than the demand-only model, except for the influence of subjective priors in the allocation model. This is different from our pricing example, where the data-generating price elasticity could not be reconciled with profit maximization, and marginal costs were partially observed.

Overall, the proposed test statistic correctly identifies the conditions where the supply side is a sensible addition to the model through improvements in the marginal conditional fit of the demand data relative to the demand-only model. Both simulation studies demonstrate that our test theory and its implementation work. The indirect Gelfand and Dey (1994) estimator of the proposed test statistic $\pi(y \mid x, M_f)$ identifies the true data-generating model from among a variety of competitive models. We find some evidence for bias in the direct NR estimator and recommend the direct estimator only as a quick first check.

5. Empirical Application

5.1. Data and Models

We illustrate our method 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. One of the outputs generated by these inputs is the net number of new customers per year, the dependent variable in our analysis. Figure 1 displays the data used in the analysis in logarithmic form. New customers for each of the regions (y) are plotted against promotional expenditures (x_1) and the number of branch offices (x_2) along with the population in each region. The points in each of the plots fall along the 45° line, corresponding to a unit slope, or a proportional allocation rule. It appears that promotional expenditures and the number of branch offices vary in direct proportion to the population in each region.

The goal is to determine whether the data are consistent with the optimal allocation rule implied by Equation (10). 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 smaller market areas satiating more quickly. We therefore standardize the variables in the analysis per a population of 1,000 and view each (standardized) geographic region as providing exchangeable information for estimating model parameters.

We specify and test alternative versions of a multiplicative production function with an additive error:

$$y_i = \gamma_i x_{1,i}^{\beta_1} x_{2,i}^{\beta_2} + \varepsilon_i, \tag{21}$$

where γ_i is a region-specific intercept that is identified through the supply-side portion of the model. Table 6 reports coefficient estimates for four analyses based on this demand model. When a supply side is present, we use the same priors for our supply-side error variance as in the simulation examples with a diffuse prior; i.e., $\pi(V_\zeta) = \mathrm{IW}(2, 0.5 \times 10^{-3} I)$. Standard methods were used for drawing parameters from their posteriors using MCMC methods, and full details are available from the authors upon request.

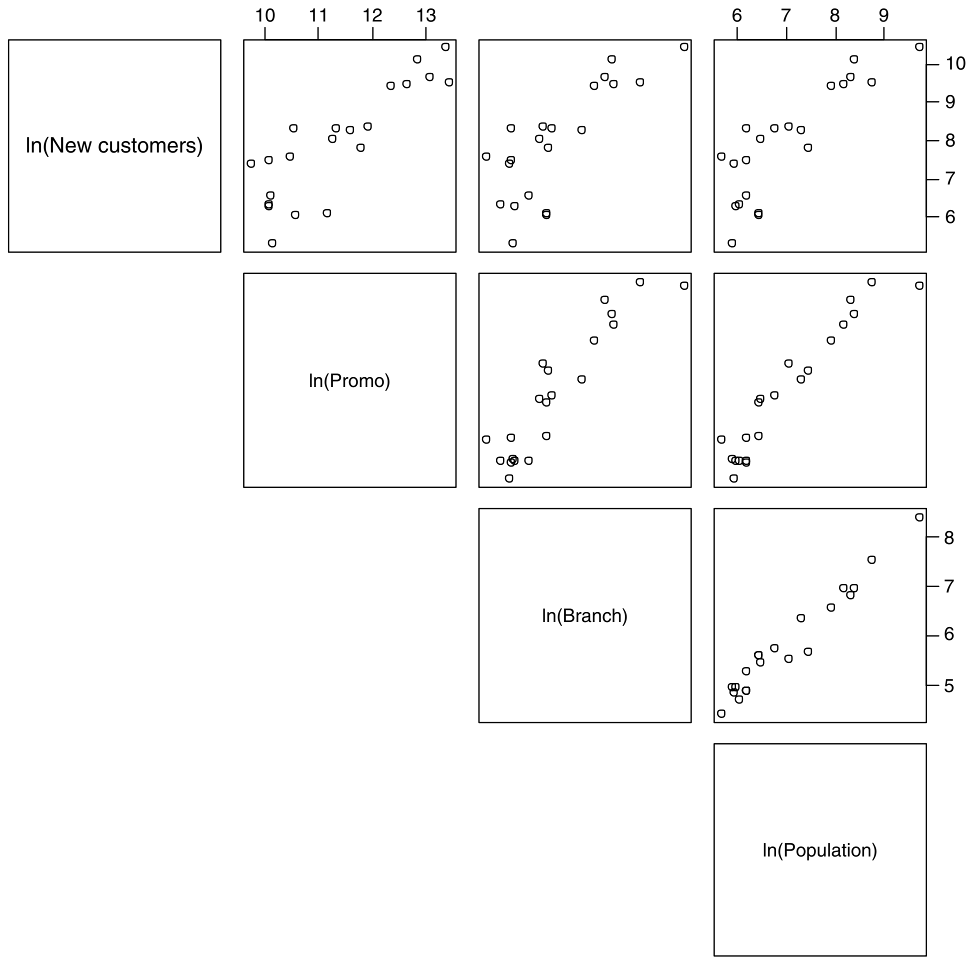
The first model contains the demand Equation (21) only. As a result, it cannot identify region-specific intercepts, although a common intercept is still identified. We find that β_2, the coefficient for the number of branch offices, is estimated to be significantly negative, which indicates that this input hinders the "production" of new customers.

The second model attempts to address the negative estimate by imposing constraints on the parameters so that $\beta_1 > 0$ and $\beta_2 > 0$, consistent with increasing returns to inputs. This constraint results in an estimate of β_2 close to zero and a slight improvement in the marginal density estimate.[5] It is important to

[4] With decreasing allocation errors, the demand-only model approaches an indeterminate design matrix, i.e., no variation in the inputs across the decision units. In this situation, adding the functional form of the allocation model to the demand model results in a noticeable gain in efficiency.

[5] We only report estimates based on GD here because we found the NR estimates to be biased in our simulation study.

Figure 1 Scatterplots of Unnormalized Data for Each Region



remember that in Bayesian analyses, the likelihood is not being maximized, and incorporating an ordinal constraint need not lead to a degradation of model fit. In fact, the additional information made present through the prior is seen to reduce uncertainty in the posterior distribution of model parameters, with the marginal density indicating a more plausible model structure.

The third model incorporates the supply-side portion of the model as a formal means for justifying the ordinal constraint, and we retain the presence of a fixed intercept for the 21 regions. We find

Table 6 Parameter Estimates: Posterior Mean

Parameters	Demand only	Demand only with constraints	Demand + Supply fixed effects	Demand + Supply random effects
β_1	0.06 (0.259)	0.14 (0.097)	0.22 (0.115)	0.20 (0.028)
β_2	−0.770 (0.370)	0.03 (0.028)	0.09 (0.109)	0.13 (0.056)
$E[\gamma_i]$	11.75 (12.07)	1.91 (0.678)	1.35 (0.647)	1.10 (0.211)
$\mathrm{Var}[\gamma_i]$	NA	NA	NA	0.21 (0.066)
$\ln \lambda_1$	NA	NA	−4.52 (0.540)	−4.56 (0.211)
$\ln \lambda_2$	NA	NA	−3.13 (1.031)	−2.38 (0.375)
$\mathrm{Var}[\zeta_1]$	NA	NA	0.08 (0.034)	0.05 (0.031)
$\mathrm{Cov}[\zeta_1, \zeta_2]$	NA	NA	−0.04 (0.020)	−0.05 (0.012)
$\mathrm{Var}[\zeta_2]$	NA	NA	0.07 (0.026)	0.08 (0.046)
σ_ε^2	3.95 (1.228)	3.86 (1.154)	3.84 (1.156)	3.99 (1.202)
Marginal density $\ln \pi(y \mid x, M)$ GD	−160.61	−158.40	−144.70	−141.23

Note. Posterior standard deviation.

an improvement of the marginal density of the conditional demand, our proposed test statistic. This indicates that the supply side implying marginally decreasing returns is supported by the demand data.

The fourth model allows for random intercepts, in other words, the presence of unobserved factors as in Equation (5) that lead to local departures from a homogeneous production function. We find that the presence of random effects identified through the supply side again improves the marginal fit of the conditional model $\pi(y \mid x, M)$. These results indicate that the firm currently follows an optimal allocation rule, and the improvements from optimal reallocations are likely to be small.

It is not obvious how this conclusion could claim empirical support without the proposed test statistic and, of course, without access to additional data. We view the ability to test whether random coefficients structurally identified through the supply side are supported by the marginal conditional fit of the demand side as one of our key contributions. We showed previously that the marginal conditional fit of the demand rejects random coefficients identified through the supply side in situations where the data were generated differently.

5.2. Optimal Allocation

Our analysis of the data indicates that our observed firm is currently allocating inputs in a manner consistent with the first-order conditions associated with optimal allocation, with departures from optimality reflected in the supply size error realizations $\zeta_{1,i}$ and $\zeta_{2,i}$. We investigate the change in expected demand through a reallocation of promotional expenditure, subject to the constraint that the total expenditure across regions and the distribution of branch offices remain the same. If the true values of the parameters β_1 and β_2 were known, then the optimal allocation of promotional expenditure must satisfy the first-order conditions:

$$\frac{\partial y_i}{\partial x_{1,i}} = \gamma_i \beta_1 x_{1,i}^{\beta_1 - 1} x_{2,i}^{\beta_2} = \gamma_k \beta_1 x_{1,k}^{\beta_1 - 1} x_{2,k}^{\beta_2} = \frac{\partial y_k}{\partial x_{1,k}} \tag{22}$$

or

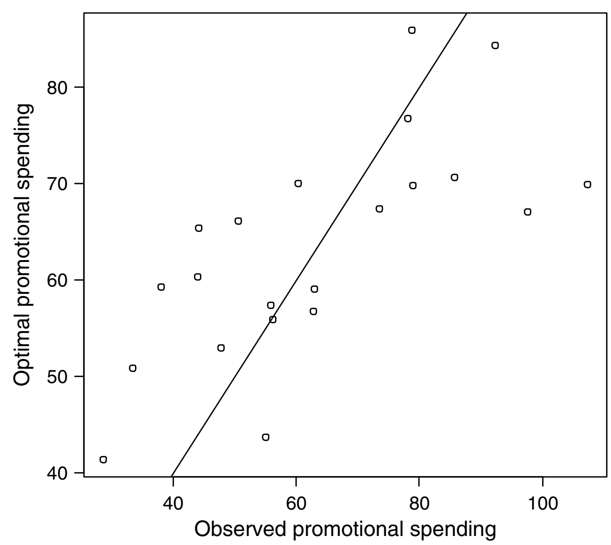
$$\frac{x_{1,i}}{x_{1,k}} = \left(\frac{\gamma_i x_{2,i}}{\gamma_k x_{2,k}}\right)^{\beta_2/(1-\beta_1)}. \tag{23}$$

Because we do not directly observe the true values of the parameters, we employ the posterior distribution of the model parameters to obtain the expected output for a given allocation of promotional expenditures and employ a search algorithm to identify the optimal allocation. Details of our algorithm are provided in Appendix B. Results of our search, applied to each of the models investigated, are displayed in Table 7.

Table 7 Optimal Allocation of Promotional Expenditures

Parameters	Demand only	Demand only with constraints	Demand + Supply fixed effects	Demand + Supply random effects
Expected y | Observed x	67.4	69.88	70.36	68.66
Expected y | Optimal x	496.38	70.34	71.06	68.98

Figure 2 Optimal and Observed Promotional Spending for the Demand + Supply Model with Random Effects



We find similar results among models except for the first model, where the estimate for β_2 is negative. For this response function, the optimal allocation is to concentrate expenditure in areas with a relatively small number of branches per 1,000 so that the negative effect is minimized. The dramatic increase in the expected new customers is due to the large estimated intercept in this model. For the remaining models, the expected increase in output (new customers per a population of 1,000) is slightly greater than expected with the current allocation.

Figure 2 displays changes in the regional allocation of promotional expenditures for the "Demand + Supply" model with random effects. The horizontal axis of the graph corresponds to the current allocation, and the vertical axis corresponds to the optimal allocation. Also plotted is a 45° line. The plot indicates that current allocation is too diverse—i.e., the optimal allocation is less extreme. However, as anticipated previously, the achieved output is already close to optimal.

6. Conclusion and Future Research

This paper introduces methods for developing and testing models of strategic behavior where the joint likelihood of demand (y) and supply (x) can be factored into a conditional factor of demand given supply, as well as a marginal factor of supply. These models of strategic behavior assume that inputs are set before outputs are realized, and they differ from models of simultaneity where inputs and outputs are inextricably jointly realized in the marketplace. Strategic models allow for a factorization of the joint likelihood of inputs and outputs (conditioning on the information decision makers might have used), for which we develop a test statistic based on the marginal density of the conditional demand data. The idea behind the test is that if the supply-side model is correct, it will help predict demand when parameter estimates are based on the joint versus the conditional likelihood. We examine the performance of the proposed test statistic in simulated and actual data, and we investigate improvements in expected output for a firm that is found to already be allocating resources in an approximately optimal manner. Without our proposed test statistic, this final conclusion would have to be based on prior knowledge or additional data.

Models of strategic behavior provide a rationale for constraining parameters so that models have reasonable implications. Inputs such as prices, promotional expenditures, and branch offices are rarely concentrated at extreme values in the marketplace, which reflects management's belief in diminishing marginal returns. The development of a supply-side specification based on the same parameters in the demand model requires the parameters to explain both input and output variables. Our proposed test statistic can be used to develop models that fit both supply and demand data without sacrificing the fit to the conditional demand data. This is especially useful in small samples, where coefficients with unreasonable values are often found.

The framework and analysis presented in this paper can be extended in a number of ways. First, although our discussion touched on industrial organization-like models, we did not explicitly apply our test statistic to such a model. An interesting avenue for future research is to determine whether the game-theoretic supply-side specifications of these models are consistent with the demand-side likelihoods. The empirical organization literature started with sparse data problems, where only the combination of demand and supply theory provided enough prior structure to identify parameters of interest at all (Reiss and Wolak 2007). With sparse data, the scope for statistically testing the assumptions needed for identification is obviously limited. However, the empirical organization literature-inspired applications in marketing, a field relatively richer in data, may benefit from the proposed test statistic. For example, using Berry's (1994) general notation, market shares are a function of observed product characteristics x, prices p, unobserved product characteristics ξ, and unobserved demand parameters θ:

$$\mathbf{s}_t = f(\mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta}). \tag{24}$$

Prices, in turn, are an implicit function of the same arguments and marginal costs, where the implicit function reflects optimization behavior in a competitive environment and defines equilibrium prices:

$$\mathbf{p}_t = g(\mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta}, \mathbf{mc}_t). \tag{25}$$

Given unobserved parameters and unobserved product characteristics, we can factor the joint density of shares and prices as follows:

$$\begin{aligned}&\pi(\mathbf{p}_t, \mathbf{s}_t \mid \mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta}, \mathbf{mc}_t)\\ &\quad= h[\mathbf{s}_t, f(\mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta})] \times k[\mathbf{p}_t, g(\mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta}, \mathbf{mc}_t)]. \end{aligned}\tag{26}$$

The function $h[\mathbf{s}_t, f(\mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta})]$ is the conditional likelihood of observed market shares $\mathbf{s}_t$ given those predicted, i.e., $f(\mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta})$, where $\mathbf{p}_t$ are observed prices. The function $k[\mathbf{p}_t, \mathbf{p}_t^* = g(\mathbf{x}, \mathbf{p}_t, \xi_t, \boldsymbol{\theta}, \mathbf{mc}_t)]$ is the marginal likelihood of observed prices given the predicted prices $\mathbf{p}_t^*$ that result from competitive optimization given knowledge about demand primitives and marginal cost. Because prices and observed market shares are seen to be informative about demand primitives in this model, it is valid to question whether adding the pricing story is supported by the marginal conditional fit of the demand data.

A concrete example is a Nash pricing scenario with expected market share of brand j at time t:

$$MS_{jt} = \frac{\exp(x'\beta_j + p_{jt}\beta_p + \xi_{j,t})}{\sum_{k=1}^{K} \exp(x'\beta_k + p_{kt}\beta_p + \xi_{k,t})} \tag{27}$$

and implied equilibrium prices

$$p_{jt} = mc_{jt} - \frac{1}{\beta_p(1 - MS_{jt})}. \tag{28}$$

With at least partially known variable costs, the pricing equation is highly informative about the price parameter, other demand parameters, and unobserved product characteristics, similar to the monopolistic pricing example explored in our simulation study. Observed prices and variable costs, coupled with the above-mentioned pricing rule, define an informative prior for the parameters in the share equations. Thus, the marginal contribution of the pricing story to fitting the market shares can be assessed after defining a demand-side likelihood that measures departures from expected shares and a supply-side likelihood that measures departures from implied prices by adding error terms. The marginal density of the market shares could be used to compare the

Nash equilibrium to a demand-only model or to other models of market behavior. We leave this for future research.

Second, 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 because it is 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.

Third, we have considered only continuous strategic decisions and not discrete strategic choices by firms. Examples might include whether or not to adopt an everyday low pricing strategy (Ellickson and Misra 2008) or the decision to continue operating in a specific market. Discrete strategic choices represent corner solutions similar to the discreteness found in choice models, and they lead to inequality constraints in the supply-side specification. Additional research is needed to develop the proposed testing ideas in this context.

Appendix A. Derivation and Calculation of Conditional Marginal Likelihoods

The direct NR-style estimator of the conditional marginal likelihood is derived as follows:

$$\begin{aligned}
&\int \frac{\pi(x\mid\theta_x,\theta_{x,y})\pi(\theta_x)\pi(\theta_{x,y})\pi(\theta_y)}{\pi(y\mid x,\theta_y,\theta_{x,y})\pi(x\mid\theta_x,\theta_{x,y})\pi(\theta_x)\pi(\theta_y)\pi(\theta_{x,y})}\\
&\quad\times\frac{\pi(y\mid x,\theta_y,\theta_{x,y})\pi(x\mid\theta_x,\theta_{x,y})\pi(\theta_x)\pi(\theta_y)\pi(\theta_{x,y})}{\pi(y\mid x)\pi(x)}\,d\{\theta_x,\theta_y,\theta_{x,y}\}\\
&\quad=\int\frac{\pi(x\mid\theta_x,\theta_{x,y})\pi(\theta_x)\pi(\theta_{x,y})}{\pi(y\mid x)\pi(x)}\,d\{\theta_x,\theta_{x,y}\}\\
&\quad=\frac{1}{\pi(y\mid x)\pi(x)}\int\pi(x\mid\theta_x,\theta_{x,y})\pi(\theta_x)\pi(\theta_{x,y})\,d\{\theta_x,\theta_{x,y}\}\\
&\quad=\frac{1}{\pi(y\mid x)\pi(x)}\pi(x)\\
&\quad=\frac{1}{\pi(y\mid x)}.
\end{aligned}\tag{A1}$$

We note that the only departure from the NR estimator of the joint marginal likelihood is the inclusion of the supply-side likelihood $\pi(x\mid\theta_x,\theta_{x,y})$ as part of the subsequently non-normalized importance density.

Next we show analytically why estimating joint marginal likelihoods using the Newton and Raftery (1994) (NR), approximation in the presence of a supply-side model is usually flawed. Then we detail how to estimate conditional marginal likelihoods based on Gelfand and Dey (1994) (GD) in our applications.

In Bayesian analysis of joint models of demand and supply, e.g., Yang et al. (2003), it is current practice to couple the joint likelihood from Equation (1) with standard priors such as in

$$\pi(\theta_x,\theta_y,\theta_{x,y}\mid y,x)\\
\propto\pi(y\mid x,\theta_y,\theta_{x,y})\times\pi(x\mid\theta_x,\theta_{x,y})\times\pi(\theta_x,\theta_{x,y},\theta_y).\tag{A2}$$

By standard priors, we mean priors that do not reflect the effective constraints on parameters through the supply-side likelihood $\pi(x\mid\theta_x,\theta_{x,y})$. Examples of such implied constraints are that own-price elasticities have to be smaller than -1 or that the sum of exponents in Cobb-Douglas production functions is constrained to values below 1. If the prior $\pi(\theta_x,\theta_{x,y},\theta_y)$ does not reflect these constraints, then the identity on which the NR estimator is based is violated a priori. That is,

$$\int\frac{\pi(\theta_x,\theta_{x,y},\theta_y)}{\pi(\theta_x,\theta_{x,y},\theta_y\mid x,y)}\\
\times\pi(\theta_x,\theta_{x,y},\theta_y\mid x,y)\,d\{\theta_x,\theta_y,\theta_{x,y}\}<1\tag{A3}$$

because the posterior does not have the same support as the prior distribution. Lenk (2009) pointed out that the NR estimator is not simulation-consistent for all practical purposes because of the large discrepancy between the *effective* support of the posterior MCMC sample and the *theoretical* support of the prior even if the *theoretical* support of the posterior is identical to that of the prior. Here, the situation is potentially worse because usually joint models of demand and supply result in a posterior that is not even *theoretically* supported over the range of standard priors. Thus, a model comparison based on NR between a structural model with constraints through the supply side and a descriptive model of y and x without implied constraints, each coupled with standard priors, is likely to favor the structural model by construction of the NR estimator.

Calculating the Indirect GD Estimator for the Monopolist Pricing Example

In our general notation, the application of Gelfand and Dey's (1994) identity requires the specification of normalized importance densities $q(\theta_x,\theta_{x,y},\theta_y)$, $q(\theta_x,\theta_{x,y})$, and $q(\theta_{x,y},\theta_y)$, with thin tails relative to $\pi(\theta_x,\theta_{x,y},\theta_y\mid x,y)$, $\pi(\theta_x,\theta_{x,y}\mid x)$, and $\pi(\theta_{x,y},\theta_y\mid y)$, i.e., the posterior from the joint model, from the supply-side-only model, and from the demand-side-only (the conditional) model, respectively. We can then apply Equation (20) to obtain the marginal likelihoods $\pi(y,x\mid M_f)$, $\pi(x\mid M_f)$, and $\pi(y\mid M_c)$, as well as the target quantity $\pi(y\mid x,M_f)$ as $\pi(y,x\mid M_f)/\pi(x\mid M_f)$, which can be directly compared with $\pi(y\mid M_c)$ and across specifications.

We will start with how to specify $q(\theta_x,\theta_{x,y},\theta_y)$ based on an MCMC sample from $\pi(\theta_x,\theta_{x,y},\theta_y\mid x,y)$. In our monopolist pricing example, we have three joint models of demand and pricing: (i) a model that rationalizes price variation through random cost shocks only, with posterior $\pi(\beta_0,\beta_1,\sigma_\varepsilon^2,\sigma_v^2\mid\{p_t,y_t\})$; (ii) a model that connects demand and price variation through common shocks, with posterior $\pi(\beta_0,\beta_1,\sigma_\varepsilon^2,\sigma_v^2,\sigma_\eta^2,\{\eta_t\}\mid\{p_t,y_t\})$; and (iii) a model where time-varying price elasticity links demand and price variation, with posterior $\pi(\beta_0,\{\beta_{1,t}\},V_{\beta1},\sigma_\varepsilon^2,\sigma_v^2\mid\{p_t,y_t\})$.

We specify importance densities as multivariate normal distributions. The exponential tails of the multivariate normal are useful because we need thin tails. However, the variance parameters σ_ε^2, σ_v^2, σ_η^2, and $V_{\beta1}$ can only take positive values, and β_1 and $\{\beta_{1,t}\}$ are restricted to values smaller than -1. Therefore we log-transform σ_ε^2, σ_v^2, σ_η^2, and $V_{\beta1}$. To obtain unrestricted counterparts of β_1 and $\{\beta_{1,t}\}$, we use the transformation $\log(-(\beta_1+1))$. We fit a multivariate normal distribution to the posterior after transformation. The importance density is then the multivariate normal density of the transformed posterior times a Jacobian. The resulting Jacobian matrix is diagonal, with entries of the form $1/V$ for variance terms and $-1/(\beta_1+1)$ for price elasticities.

The specification of $q(\beta_1, \sigma_v^2)$, $q(\beta_1, \sigma_v^2, \sigma_\eta^2, \{\eta_t\})$, and $q(\beta_1, \{\beta_{1,t}\}, V_{\beta 1}, \sigma_v^2)$ based on MCMC samples from $\pi(\beta_1, \sigma_v^2 \mid \{p_t\})$, $\pi(\beta_1, \sigma_v^2, \sigma_\eta^2, \{\eta_t\} \mid \{p_t\})$, and $\pi(\beta_1, \{\beta_{1,t}\}, V_{\beta 1}, \sigma_v^2 \mid \{p_t\})$ to estimate $\pi(x \mid M_f)$ under the three specifications proceeds analogously.

In the case of the demand-only model, the three posteriors $\pi(\beta_0, \beta_1, \sigma_\varepsilon^2 \mid \{y_t\})$, $\pi(\beta_0, \beta_1, \sigma_\varepsilon^2, \sigma_\eta^2, \{\eta_t\} \mid \{y_t\})$, and $\pi(\beta_0, \{\beta_{1,t}\}, V_{\beta 1}, \sigma_\varepsilon^2 \mid \{y_t\})$ are well defined because of proper priors and functional form assumptions. However, for all practical purposes, only $\pi(y \mid M_c)$ corresponding to the model $\pi(\beta_0, \beta_1, \sigma_\varepsilon^2 \mid \{y_t\})$ is relevant. The specification of the importance density proceeds in analogue to the joint model above.

Calculating the Indirect GD Estimator for the Cobb-Douglas Example

In our Cobb-Douglas example, we have two joint models of output y and inputs x: (i) a model that rationalizes input variation through random cost shocks, input pricing errors, or both, with posterior $\pi(\beta_0, \log(\beta_1), \log(\beta_2), \sigma_\varepsilon^2, \Sigma_\zeta, \{\zeta_i\} \mid \{x_{1,i}, x_{2,i}, y_i\})$; and (ii) a model that connects variation in the inputs and output across units to unobserved differences in the production function, with posterior $\pi(\{\beta_{0,i}\}, V_{\beta 0}, \log(\beta_1), \log(\beta_2), \sigma_\varepsilon^2, \Sigma_\zeta, \{\zeta_i\} \mid \{x_{1,i}, x_{2,i}, y_i\})$.

We follow a similar strategy to specify importance densities based on MCMC samples from the posterior as in our monopolist pricing example. However, in the Cobb-Douglas case, we have a constraint on the exponents in the production function, i.e., $\beta_1 + \beta_2 < 1$, which is impossible to represent by fixed transformations. The importance density needs to reflect this constraint coming through the supply side. If it does not, it will have a fatter tail than the posterior in the direction of $\beta_1 + \beta_2 >= 1$ by definition, and the GD identity from Equation (20) will no longer hold. We view this and similar complications resulting from implicit constraints as a characteristic features of models that are derived from the optimization behavior of agents.

Our solution to this problem first factorizes the joint importance density $q(\beta_0, \log(\beta_1), \log(\beta_2), \sigma_\varepsilon^2, \Sigma_\zeta, \{\zeta_i\})$ into a marginal part corresponding to the constraint with no explicit representation $q(\log(\beta_1), \log(\beta_2))$ and then a conditional part where all constraints, if any, can be represented by simple transformations $q(\beta_0, \sigma_\varepsilon^2, \Sigma_\zeta, \{\zeta_i\} \mid \log(\beta_1), \log(\beta_2))$. We define $q(\log(\beta_1), \log(\beta_2))$ as a bivariate normal distribution fitted to $\pi(\log(\beta_1), \log(\beta_2) \mid \{x_{1,i}, x_{2,i}, y_i\})$ truncated to the range where $\beta_1 + \beta_2 < 1$. We estimate the mean and variance–covariance of the corresponding unconstrained bivariate normal as well as the normalizing constant by simulation.

We then define $q(\beta_0, \sigma_\varepsilon^2, \Sigma_\zeta, \{\zeta_i\} \mid \log(\beta_1), \log(\beta_2))$ as $q(\beta_0, \sigma_\varepsilon^2, \{\zeta_i\} \mid \log(\beta_1), \log(\beta_2)) \times q(\Sigma_\zeta)$. We define $q(\Sigma_\zeta)$ independently for convenience as an Inverse Wishart density with parameters estimated from $\pi(\Sigma_\zeta \mid \{x_{1,i}, x_{2,i}, y_i\})$.

The density $q(\beta_0, \sigma_\varepsilon^2, \{\zeta_i\} \mid \log(\beta_1), \log(\beta_2))$ is specified as $\text{MVN}(\beta_0, \log(\sigma_\varepsilon^2), \{\zeta_i\} \mid \log(\beta_1), \log(\beta_2)) \times 1/\sigma_\varepsilon^2$. We estimate the means and the variance–covariance matrix of these distributions, regressing $\beta_0, \log(\sigma_\varepsilon^2), \{\zeta_i\}$ on $\log(\beta_1), \log(\beta_2)$ in the MCMC sample from $\pi(\beta_0, \log(\beta_1), \log(\beta_2), \sigma_\varepsilon^2\{\zeta_i\} \mid \{x_{1,i}, x'_{2,i} y_i\})$.

Estimation of $\pi(x \mid M_f)$ under the two specifications does not present any additional difficulties given samples from $\pi(\beta_0, \log(\beta_1), \log(\beta_2), \Sigma_\zeta, \{\zeta_i\} \mid \{x_{1,i}, x_{2,i}\})$, and $\pi(\{\beta_{0,i}\}, V_{\beta 0}, \log(\beta_1), \log(\beta_2), \Sigma_\zeta, \{\zeta_i\} \mid \{x_{1,i}, x_{2,i}\})$. This is similar to estimating $\pi(y \mid M_c)$. Again, even though $\pi(\{\beta_{0,i}\}, V_{\beta 0}, \log(\beta_1), \log(\beta_2), \sigma_\varepsilon^2 \mid \{y_i\})$ is well defined with proper priors and our functional form assumptions, the comparison in practice will be to a demand model without random effects, i.e., $\pi(\beta_0, \log(\beta_1), \log(\beta_2), \sigma_\varepsilon^2 \mid \{y_i\})$, when only a cross section is observed.

Appendix B. Optimal Actions

In general, the problem of finding the optimal Bayes action is to find $a^* = \max_a \int U(a, \theta) p(\theta \mid data)\, d\theta$, where $U(a, \theta)$ is the utility to be maximized. It depends on the action a and characteristics of nature θ. The optimal action is defined as the value of the action that maximizes (posterior) expected utility.

In our case, utility is equal to the (expected) output y as a function of the actions or inputs $\{x_{1i}, x_{2i}\}$ and characteristics of the areas producing the outputs summarized by β_1, β_2, and γ (or $\{\gamma_i\}$, depending on the model), $y(\{x_{1i}, x_{2i}\}, \beta_1, \beta_2, \{\gamma_i\})$. If the response parameters were known, we could obtain closed-form solutions for the optimal inputs by numerically solving (23) for the inputs, subject to constraints. Because our knowledge of response parameters is limited by posterior uncertainty, the optimal allocation is defined as the maximum of an integral. In general, we have

$$\begin{aligned} \{x^*_{1,i}, x^*_{2,i}\} = \max_{\{x_{1,i}, x_{2,i}\}} & \int \sum_{i=1}^{N} y_i(x_{1,i}, x_{2,i}, \beta_1, \beta_2, \gamma_i) \\ & \cdot p(\{\beta_1, \beta_2, \gamma_i\} \mid data, M)\, d\{\beta_1, \beta_2, \gamma_i\} \\ \text{s.t.} \quad & c(\{x_{1,i}, x_{2,i}\}) \le \text{constraint}. \end{aligned} \tag{B1}$$

In our application, we lack the price information needed to optimally balance the two inputs $x_{1,i}$ and $x_{2,i}$ within areas. Thus, we investigate the conditionally optimal allocation of the first input across areas. Note that we implicitly assume that the price of the first input is constant across areas. Thus we solve the following maximization problem

$$\begin{aligned} \{x^*_{1,i}\} = \max_{\{x_{1,i}\}} & \int \sum_{i=1}^{N} y_i(x_{1,i}, x_{2,i}^{\text{observed}}, \beta_1, \beta_2, \gamma_i) \\ & \cdot p(\{\beta_1, \beta_2, \gamma_i\} \mid data, M)\, d\{\beta_1, \beta_2, \gamma_i\} \\ \text{s.t.} \quad & \sum_{i=1}^{N} x_{1,i} = \sum_{i=1}^{N} x_{1,i}^{\text{observed}}. \end{aligned} \tag{B2}$$

Because our production function is nonlinear and the posterior distribution is not of a known form, a closed-form solution to this maximization problem is not available. We use a simulated annealing algorithm to approximate the conditionally optimal inputs. The algorithm picks two areas i and j at random, increases (decreases) $x_{1,i}$ by a randomly determined delta, and decreases (increases) $x_{1,j}$ by the same amount. The candidate reallocation is accepted with probability

$$\min\left(\exp\left(\kappa\left[\sum_{i=1}^{N} y_i(x_{1,i}^{\text{cand}}, p(\{\beta_1, \beta_2, \gamma_i\} \mid data, M)) - \sum_{i=1}^{N} y_i(x'_{1,i} p(\{\beta_1, \beta_2, \gamma_i\} \mid data, M))\right]\right), 1\right). \tag{B3}$$

The positive scale parameter κ controls the "heat" of the algorithm. Small values of κ correspond to more heat, allowing the algorithm to traverse areas of decreasing expected output. We increase κ as a function of the iterations of the simulated annealing algorithm such that, eventually, the local maximum is determined almost surely.

References

Berry, S. T. 1994. Estimating discrete-choice models of product differentiation. *RAND J. Econom.* **25**(2) 242–262.

Bronnenberg, B. J., V. Mahajan. 2001. Unobserved retailer behavior in multimarket data: Joint spatial dependence in market shares and promotion variables. *Marketing Sci.* **20**(3) 284–299.

Carroll, J. D., P. E. Green, W. S. DeSarbo. 1979. Optimizing the allocation of a fixed resource: A simple model and its experimental test. *J. Marketing* **43**(1) 51–57.

Chamberlain, G. 1984. Panel data. Z. Griliches, M. D. Intrilligator, eds. *Handbook of Econometrics*, Vol. 2. North Holland, Amsterdam, 1247–1318.

Chintagunta, P. K., T. Erdem, P. E. Rossi, M. Wedel. 2006. Structural modeling in marketing: Review and assessment. *Marketing Sci.* **25**(6) 604–616.

Dong, X., P. Manchanda, P. K. Chintagunta. 2009. Quantifying the benefits of individual-level targeting in the presence of firm strategic behavior. *J. Marketing Res.* **46**(2) 207–221.

Ellickson, P. B., S. Misra. 2008. Supermarket pricing strategies. *Marketing Sci.* **27**(5) 811–828.

Gelfand, A. E., D. K. Dey. 1994. Bayesian model choice: Asymptotics and exact calculations. *J. Roy. Statist. Soc. Ser. B* **56**(3) 501–514.

Horsky, D., P. Nelson. 1996. Evaluation of salesforce size and productivity through efficient frontier benchmarking. *Marketing Sci.* **15**(4) 301–320.

Lenk, P. 2009. Simulation pseudo-bias correction to the harmonic mean estimator of integrated likelihoods. *J. Comput. Graphical Statist.* **18**(4) 941–960.

Manchanda, P., P. E. Rossi, P. K. Chintagunta. 2004. Response modeling with nonrandom marketing-mix variables. *J. Marketing Res.* **41**(4) 467–478.

Mantrala, M. K., P. Shinha, A. A. Zoltners. 1992. Impact of resource allocation rules on marketing investment-level decisions and profitablility. *J. Marketing Res.* **29**(2) 162–175.

Misra, S., H. Nair. 2009. A structural model of sales-force compensation dynamics: Estimation and field implementation. Research Paper 2037, Stanford University, Stanford, CA. http://ssrn.com/abstract=1474462.

Misra, S., A. T. Coughlan, C. Narasimham. 2005. Salesforce compensation: An analytical and empirical examination of the agency theoretic approach. *Quant. Marketing Econom.* **3**(1) 5–39.

Morey, R. C., J. M. McCann. 1983. Estimating the confidence interval for the optimal marketing mix: An application to lead generation. *Marketing Sci.* **2**(2) 193–202.

Newton, M. A., A. E. Raftery. 1994. Approximate Bayesian inference with the weighted likelihood bootstrap. *J. Roy. Statist. Soc. Ser. B* **56**(1) 3–48.

Reiss, P. C., F. A. Wolak. 2007. Structural econometric modeling: Rationales and examples from industrial organization. J. J. Heckman, E. E. Leamer, eds. *Handbook of Econometrics*, Vol. 6A. Elsevier, Amsterdam, 4277–4415.

Rossi, P. E., G. M. Allenby, R. McCulloch. 2005. *Bayesian Statistics and Marketing*. John Wiley & Sons, New York.

Villas-Boas, J. M., R. S. Winer. 1999. Endogeneity in brand choice models. *Management Sci.* **45**(10) 1324–1238.

Yang, S., Y. Chen, G. M. Allenby. 2003. Bayesian analysis of simultaneous demand and supply. *Quant. Marketing Econom.* **1**(3) 251–276.

Zellner, A., J. Kmenta, J. Drèze. 1966. Specification and estimation of Cobb-Douglas production function models. *Econometrica* **34**(4) 784–795.