DISCUSSION PAPER

April 2007, revised July 2007 ■ RFF DP 07-25-REV

Non-Price Equilibria for Non-Marketed Goods

Daniel J. Phaneuf, Jared C. Carbone,
and Joseph A. Herriges

The Frontiers of Environmental Economics

Hosted by Resources for the Future
1616 P Street, NW
Washington, DC 20036
202.328.5000
www.rff.org/frontiersconference

NON-PRICE EQUILIBRIA FOR NON-MARKETED GOODS*

by

Daniel J. Phaneuf, Jared C. Carbone, and Joseph A. Herriges

Commissioned for: Resources for the Future *Frontiers of Environmental Economics Workshop*,

Washington, DC February 27-28, 2007

Version: June 28, 2007

---

* The authors are Associate Professor of Agricultural and Resource Economics, North Carolina State University, Assistant Professor of Economics, Williams College, and Professor of Economics, Iowa State University, respectively. We thank without implicating Kerry Smith, Roger von Haefen, Chris Timmins, and several workshop participants for valuable discussions and comments on aspects of this project. Funding for the IA Lakes Project data used in this analysis was provided by the Iowa Department of Natural Resources and the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program. Although the research described in the article has been funded in part by the U.S. Environmental Protection Agency's STAR program through grant R830818, it has not been subjected to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred

Send correspondence to: Prof. Daniel Phaneuf, NCSU Box 8109, Raleigh, NC 277695, (919) 515-4672, dan_phaneuf@ncsu.edu.

Non-Price Equilibria for Non-Marketed Goods

Abstract:

As part of the Resources for the Future *Frontiers of Environmental Economics* collection of papers, we consider the problem of general equilibrium feedback effects in non-price space as they relate to non-market valuation. Our overall objective is to examine the extent to which non-price equilibria arising from both simple and complex sorting behavior can be empirically modeled and the resulting differences in partial and general equilibrium welfare measures quantified. After motivating the problem in general we consider the specific context of congestion in recreation demand applications, which we classify as the outcome of a simple sorting equilibrium. Using both econometric and computable general equilibrium (CGE) models we examine the conceptual and computational challenges associated with this class of problems and present findings on promising solution avenues. We demonstrate the relevance of accounting for congestion effects in recreation demand with an application to lake visits in Iowa. Our econometric and CGE results confirm that, for some plausible counterfactual scenarios, substantial differences exist between partial and general equilibrium welfare estimates. We conclude the paper by describing tasks that are needed to move forward research in this area.

JEL Classification: C35, D58, Q25, Q26

I. Introduction

Empirical studies of market activities draw on an elegant and coherent body of theory that describes household and firm interactions in the market place. Price taking households purchase goods produced by firms that compete to maximize profits under a variety of market power conditions. Theory provides behavioral predictions for households and firms as well as statements about how the aggregation of this behavior results in equilibrium price and quantity outcomes. Models of general equilibrium rely on this link between individual behavior and aggregate outcomes to describe how exogenous changes lead to both direct and indirect effects in price and quantity space. Often it is the indirect or feedback, effects, that are the most interesting in market studies. A variety of empirical and calibration techniques have been developed in economics to study these effects. The modern empirical IO literature focusing on particular industries provides a good example of the former while CGE models of whole sectors of the economy provide good examples of the latter. In both cases the emphasis is on modeling and understanding equilibrium outcomes in price and quantity space.

The story is quite different in studies of non-market goods that are typically employed by environmental economists for purposes of non-market valuation. By definition non-marketed goods are not exchanged in markets, and therefore one cannot speak of equilibrium prices and quantities for the goods per se. Instead the emphasis is usually on understanding preferences in a partial equilibrium framework for a quasi-fixed level of a public good. For this purpose an impressive array of structural econometric models has been developed that are capable of predicting individuals' valuations for exogenous changes in the level of the public good. For example, recreation demand modelers use increasingly sophisticated models of quality differentiated demands to understand how recreation site attributes affect behavior and well-being. Hedonic property value models use ever increasing levels of spatially resolute data to parse out the contribution of a local public good to housing prices. Although the latter models make use of equilibrium concepts to motivate estimation, they rarely are capable of predicting feedback effects from large-scale changes in public good levels. Thus, with few exceptions, it seems reasonable to say that non-market valuation has focused primarily on partial equilibrium analysis of the interactions between behavior and quasi-fixed levels of environmental quality.[1]

This emphasis is probably reasonable in general. Empirical models of behavior that use measurable environmental quality as explanatory variables usually find effects that are of second order importance relative to non-environmental factors. For example, ambient water quality in recreation demand models is usually much less important in explaining site choice and visitation frequency than travel cost. Likewise, structural characteristics tend to explain much more of the variability in housing prices than does air quality in hedonic property value models. Water quality and air quality in these contexts are examples of non-price attributes that we might reasonably suppose to be exogenous to the behavior that we are attaching to them. In contrast, the levels of other types of attributes – such as congestion or angler catch rates in recreation models, or traffic levels in residential location choice models – are at least partially determined by the aggregation of behavior under analysis. We may therefore wonder if there are situations in which general equilibrium feedback effects in endogenous attribute space might be empirically important in non-market valuation. This might be particularly so for a large-scale policy intervention that substantially changes the level and spatial distribution of environmental quality. In this paper we begin to consider the extent to which non-price equilibria and feedback effects can be identified and accounted for conceptually and in empirical non-market valuation studies.

To examine this question we proceed as follows. We begin by providing a descriptive overview of how we will think about the concept of 'non-price equilibria' in non-market valuation. We suggest working definitions of two general types of non-price equilibria and offer context and motivation by linking these definitions to specific examples and the existing literature. We then turn to study a specific type of non-price endogenous attribute: congestion in recreation demand models. We do this using both econometric and computable general equilibrium (CGE) models. We begin by laying out a general modeling structure that will be used in both the econometric and CGE models, which is followed by a description of the Iowa lakes data that motivates our empirical analysis in both modeling frameworks. We then consider an empirical model of recreation demand that explicitly includes site congestion as an explanatory variable, accounts for its econometric endogeneity, and allows computation of both partial and general equilibrium welfare measures. After presenting results from this analysis we turn to the CGE model, which is used to explore more generally situations when partial and general equilibrium welfare measures can be different and under what circumstances it might be important to consider non-price feedback effects in empirical non-market valuation models.

With the three components of this paper we provide three contributions in the spirit of the 'frontiers' theme of this conference. First, we lay out a research agenda that is motivated by the notion that large scale policy interventions might lead to feedback effects in non-price variables. Second, we make use of both CGE and econometric modeling approaches to analyze the feedbacks problem and demonstrate how these quite different tools can shed light on the same problem from different angles. Since the behavior we are interested in is characterized by both intensive and extensive margin decisions both modeling approaches must admit binding non-negativity constraints. Thus a second contribution of this work is the further development of CGE and econometric models that are flexible, tractable, and provide realistic and internally consistent representations of the behavior we are modeling. The final contribution involves the application to recreation visits to Iowa lakes and accounting for congestion in the model.

II. Conceptual Overview

To ground our discussion of non-price equilibria we consider the following behavioral setup. Agents in a closed economy maximize an objective function by choosing levels of activities that are defined by both price and a set of non-price attributes. In the case of consumers the activities are demands for quality-differentiated goods; for firms we can think of them as derived demands for quality-differentiated factor inputs. For the remainder of this discussion we use terminology corresponding to the consumer’s problem, although we will also provide examples that correspond to firm’s behavior. Households consume the quality differentiated goods in non-negative quantities and can, at the optimum, be at a corner solution for a subset of the available goods. The set of non-price attributes that describe the goods in the choice set can be divided into two types: those that are exogenously determined and those that are at least partially determined by the actions of individuals in the model (i.e., endogenous attributes). We are ultimately interested in understanding the extent to which the levels of endogenous attributes might change in response to exogenous or policy shocks, and what the resulting differences are between partial and general equilibrium welfare measures for policy interventions.

This general setup can be better understood by adding a few specific examples. The most obvious case is when the quality differentiated goods are trips to recreation sites, say a set of lakes. The demand for trips depends on individual travel costs as well as attributes of the recreation sites. Attributes such as the presence of boat ramps, picnic facilities, and perhaps ambient water quality are exogenous to the decision process. In contrast, congestion at the recreation sites is determined by the aggregate visitation decisions of the people in the economy and is therefore an endogenous attribute. Similarly, angling catch rates for sport fish species at the lakes are determined not only by existing bio-physical conditions, but also by the spatial and temporal distribution of anglers' fishing effort. Policy interventions such as water quality improvements, facility improvements, or fish stocking programs might have direct welfare effects as well as indirect effects that play through via the re-equilibrating of congestion and catch rate attributes that change due to people's changed visitation patterns.

A second example is the choice of residential location, which conveys a bundle of market and non-market services. Exogenous attributes in the bundle include characteristics of the structure and distance to natural features such as lakes. Endogenous attributes might include traffic congestion and the resulting local air quality impacts, privately held open space, and publicly held open space created by local governance. This latter attribute is related to other endogenous attributes that have more of a public finance or urban economic flavor, such as local school quality and the racial makeup of neighborhoods.

A final example comes from commercial fisheries. Vessel operators choose the timing and location of harvest effort that gives rise to an aggregate distribution of effort in much the same way that congestion is determined in a recreation model. This effort, in combination with the biological system, gives rise to equilibrium populations for the targeted species as well as equilibrium levels of spatially and temporally distributed catch effort.

These examples suggest two general types of non-price equilibria that can occur in environmental economics applications. The first case we refer to as a *simple sorting equilibrium*. In this case levels of endogenous non-price attributes are determined only by interactions among agents. Among the examples mentioned this class includes congestion in recreation applications and more generally social interaction outcomes such as racial mixing and peer effects in schools. In these cases the equilibrium outcomes are determined only by the interactions among agents. This is in contrast to the second type of non-price equilibrium that we define which we label a *complex sorting equilibrium*. This refers to situations in which the agents interact with a quasi-supplier (often the natural environment) to determine the equilibrium. Among the examples we have mentioned recreation fishing catch rates fall into this class. Here the natural environment provides the stock of fish while anglers' aggregate distribution of trips and catch effort provides the level of stock exploitation. The interaction results in equilibrium catch rate levels and fish populations. Educational outcomes in public finance applications are similarly complex sources of sorting equilibrium. Here the sorting behavior of households into school districts determines peer effects, while district funding levels determine teacher and facility quality. Together these factors determine educational outcomes.

To formalize these ideas consider the following model of consumer behavior related to the choice of quality differentiated goods. There are N consumers in the economy, each of whom maximizes utility by choosing the levels of a J-dimension vector of quality differentiated goods denoted by z_i and spending on other goods x_i. The problem is given analytically by

$$\max_{z_i, x_i} U_i = U_i(z_i, x_i; Q) \quad s.t. \quad p_i' z_i + x_i \leq y_i, \quad z_{ij} \geq 0, \quad j = 1, \ldots, J, \tag{1}$$

where U_i is the utility index for person i, Q is an $M \times J$ matrix of non-price attributes where M is the number of attributes, p_i is the vector of prices for the quality differentiated goods, and y_i denotes the person's income. Some or all of the quality attributes in the model may be endogenously determined by the aggregation of consumers' behavior and aspects of the natural environment. Thus we define an *attribute transmission function* by

$$q_{mj} = q_{mj}(z_1, \ldots, z_N; E), \quad m = 1, \ldots, M, \quad j = 1, \ldots, J, \tag{2}$$

where q_{mj} is an individual element of the quality matrix Q, $z_1, \ldots, z_N$ denotes the aggregate behavior by households in the economy, and E stands for the natural environment or a quasi-supplier. The transmission function shown in (2) is general in that it describes an endogenous attribute arising from a complex sorting equilibrium, but it also nests as special cases the two additional types of attributes we have defined: (a) $q_{mj}=q_{mj}(E)$ for exogenous attributes and (b) $q_{mj}=q_{mj}(z_1,\ldots,z_N)$ for endogenous attributes arising from a simple sorting equilibrium.

In making their choices we assume individuals take levels of the attributes as given, so the first-order conditions for person i are:

$$\frac{\partial U_i(z_i,x_i;Q)/\partial z_{ij}}{\partial U_i(z_i,x_i;Q)/\partial x_i} \le p_{ij}, \quad z_{ij}\left[\frac{\partial U_i(z_i,x_i;Q)/\partial z_{ij}}{\partial U_i(z_i,x_i;Q)/\partial x_i} - p_{ij}\right] = 0, \quad z_{ij} \ge 0, \quad j=1,\ldots,J. \quad (3)$$

Equilibrium in the economy is characterized by the simultaneous solution to a system of equations and their associated complementary slackness relationships that correspond to:

- The first-order conditions described by (3) for all individuals i = 1,…,N and
- The attribute transmission functions in (2) for j=1,…,J and $m \in \tilde{M}$, where $\tilde{M} \subseteq \{1,\ldots,M\}$ is the set of endogenous attributes.

In defining non-price equilibria we have thus far tried to be fairly general, but it is clear that this can only be taken so far. Unlike price and quantity equilibrium concepts for homogenous goods, for which theory provides quite general results, non-price equilibria for quality differentiated goods are by definition context specific. The challenge for applied welfare analysis is to characterize conceptually and empirically the particulars of the equilibrium of interest. Simple sorting equilibria seem easier to deal with than their complex counterparts in that for the former the analyst need only specify the mechanism through which agents interact, while for the latter agent interactions a production function and the relationship between the two must be spelled out.

III. Literature Review

Concerns about equilibrium concepts defined over non-price attributes are not new to the economics literature, though they have typically taken a back seat to analyses based on price movements. Often these discussions are cast in the context of spatial sorting. Schelling [26] provides one of the earlier examples, illustrating qualitatively how non-market adjustments might lead to surprising equilibrium outcomes in the context of racial segregation. A second example concerns school quality and the role of peer effects in educational outcomes. These areas have been the subject of substantial research, and include recent studies that apply concepts and approaches similar to what we draw upon in this paper. Two examples are [3] and [12]. Bayer and McMillian [3] study the role of racial sorting in determining endogenous neighborhood quality levels such as average education attainment. Their hypothesis is that preferences for racial clustering may contribute to lower equilibrium neighborhood amenities for blacks, particularly for average education levels, since the proportion of college educated blacks is smaller than for whites. To test this hypothesis they estimate a sorting model of residential location choice and use the model to conduct counter-factual analysis related to neighborhood race and average educational level outcomes. They find, for example, that a hypothetical weakening of preferences among blacks to live in the vicinity of other blacks reduces the racial gap in neighborhood amenities, implying that racial preferences are nearly as important as differences in socio-economic status in explaining the observed amenity gaps. For each of their counterfactual experiments new non-price *simple sorting* equilibrium outcomes for amenities and race patterns must be calculated.

Ferreyra [12] examines the general equilibrium impacts of school voucher programs. Households have preferences for location characteristics and school quality as well as the secular status of private school options. Private and public school quality are endogenously determined in equilibrium by the composition of households within a district *and* a school quality production function that depends on spending per student. In this sense the non-price equilibria considered by [12] are *complex sorting* outcomes. Ferreyra considers the impacts different hypothetical voucher programs. She finds that private school enrollment expands and residential location choice is affected, suggesting that vouchers break the 'bundling' that typically exists in the purchase of public school quality and residential location. These findings depend critically on the ability to simulate new equilibrium outcomes using the structure of the estimated model.

Recent papers in empirical industrial organization (e.g., [21,25]) provide additional examples of explicit attention to sorting equilibrium outcomes, notably in the context of firms' entry location decisions. Geography is a factor over which firms often seek to differentiate their products. This is particularly the case for retail operations in which proximity to potential consumers increases the payoff from entry into a location. Competitive pressure from other retailers in close proximity, however, decreases the payoff. Thus firms must balance competing factors in deciding whether or not to place a store in a given location, and must make the decision based on expectations of what competing firms will decide to do. Thus simple sorting equilibria arise in these contexts and must be accounted for empirically.

There is also research in the environmental literature that addresses aspects of the non-price equilibrium concepts we consider in this paper. The most closely related study to ours is Timmins and Murdock [30], who consider the role of congestion in a site choice model of recreation demand. Congestion in their model is a (dis)amenity determined endogenously by the aggregate proportion of people who chose to visit a particular site. [30] considers both partial and general equilibrium welfare measures for changes in site characteristics and availability, where the latter takes account of agents' re-sorting (and hence new congestion levels) in response to exogenous changes. The authors find substantial differences between the two types of welfare measures, and show that congestion is a major determinant of visitors' choices. Congestion as defined in [30] falls under our simple sorting equilibrium category. A second and more complex consideration of an endogenous amenity is Walsh [33]. Using a residential sorting model [33] examines the role of public and private open space amenities in which both are endogenously determined. The author models feedback effects associated with urban land use policies, and finds several surprising results that arise based on his ability to simulate counterfactual land use outcomes in response to policy shocks. Most importantly he finds that public open space creation crowds out private open space in general equilibrium, in some cases so much that the total open space in an urban area *falls* following a public initiative.

A second area of environmental economics that is relevant for our study is the recent work on estimating integrated biological and human systems. For example, [19] examines the impact of water quality improvements on recreational fishing in Maryland's costal bays. The unique feature of their analysis is that, while they consider only exogenous water quality changes, these changes impact recreational fishing choices indirectly through a bio-economic model of the coastal fishery. The dynamic evolution of the fish stock is modeled as a function of water quality, allowing water quality improvements to directly affect fish stocks and indirectly affect anglers via improved catch rates. Additional examples in bio-economics are provided by [27,28] in commercial fisheries. These papers are distinguished by linked economic and biological models in which fish stocks are endogenously determined by the interactions between fishing behavior and the underlying biology.

Although different in their objectives these examples from the housing demand, public finance, industrial organization, and environmental literatures share similar challenges. These challenges can be categorized along three dimensions. The first involves econometrically identifying the effect of an endogenous aggregate outcome in individual behavioral models such as recreation or residential site choice. Manski [18] provides an intuitive overview of what he calls the reflection problem in the context of nonmarket interactions, which is particularly relevant for understanding simple sorting outcomes. He describes the difficulty inherent in "…inferring the nature of an interaction process from observations on its outcomes (when outcome data) typically have only limited power to distinguish among alternative plausible hypotheses" (p. 123). From a practical perspective this involves distinguishing between actual interactions, correlated unobservables, and non-random treatments across agents. For the case of congestion in recreation demand models the challenge is to separately assess how aggregate site visitation patterns reflect both the observed and unobserved site attributes that increase visitation, and congestion externalities that discourage visitation. The second challenge involves defining equilibrium concepts and describing their properties. Non-price equilibria will often be based on contexts that are specific to the problem under study. Transmission functions, equilibrating forces, and the existence and uniqueness of equilibria need to be separately specified and examined, and these tasks can present non-trivial difficulties. [6], for example, discusses the properties of equilibria that arise in discrete choice models with simple sorting outcomes that can take the form of a congestion or agglomeration effect. While they are able to establish general existence and bounded uniqueness results it is not immediately clear that these findings extend to more complicated modeling environments. The third challenge that arises in modeling non-price equilibrium outcomes is computational. In each of the studies described above the objective is to assess the effect of a counterfactual change in some exogenous factor. This involves solving the model to determine a new equilibrium – a task that will usually involve numerical methods nested in multiple layers of simulation.

The new classes of locational equilibrium sorting models represent the major empirical frameworks for examining non-price equilibria and addressing the three modeling challenges. Our work in this paper draws heavily on insights from the two major strands of this literature: vertical and horizontal modeling approaches. The former is based on Epple and Sieg [11], which has its intellectual roots in the public finance literature focused on understanding how populations stratify across urban landscapes. In this approach to modeling households agree on a ranking of choice alternatives – typically neighborhoods – based on price and non-price attributes. Assumptions on preferences are then used to derive equilibrium conditions that are assumed to hold for any observed outcome in the data. These conditions are used to recover estimates of preferences for attributes of the neighborhoods – including, perhaps, endogenous attributes as in [33]. [29], in contrast, examine the equilibrium effects of an exogenous change in air quality stemming from reduced ozone levels. Sorting in this case occurs on the basis of preferences for housing, education and air quality in the LA Air Basin between 1989 and 1991.

The horizontal strand of literature has its intellectual roots in the industrial organization literature and is based in large part on Berry [7] and Berry *et al*. [8]. These models take the familiar discrete choice framework as the basis for demand modeling and exploit a two stage estimation strategy that allows the use of common IV methods to account for endogenous choices attributes. [4] describes how these tools can be used in a residential sorting context and [5] and [30] provide examples in environmental applications. A distinguishing feature of these models is the emphasis on a single choice outcome, such as choice of a place to live or recreation site to visit. In our analysis below we use many of the modeling techniques introduced in the horizontal sorting models, but generalize the analysis to include a broader choice environment that considers both intensive and extensive margin decisions. We also pursue a Bayesian approach to modeling, examined in the context of industrial organization models by [35].

IV. Modeling Framework

In order to assess the feedback effects we are interested in, the non-price equilibrium outcomes need to be linked in a consistent manner to estimable models of individual behavior. The models must be able to capture consumer response to changing price and quality conditions and allow for those responses to occur at both the intensive and extensive margins. This is particularly important when evaluating major policy shifts that will induce some individuals to enter or leave the market entirely (i.e., when corner solutions emerge). What is required is a model that readily admits the concept of a virtual price. By virtual price we mean a summary measure that consistently and succinctly captures all constraints on behavior, both price and non-price, which in turn ultimately determines people's choices. The KT model [23,32] and its dual counterpart [17,22,24] are uniquely positioned for this purpose. The latter is particularly attractive in that it involves a direct parameterization of individuals' virtual prices.

In this section of the paper we provide a general overview of the dual modeling framework and a discussion of the particular functional form that will be used in our application and CGE analysis. A detailed discussion of the estimation procedure is delayed until section VI, which follows our description of the application in section V.

A. The Dual Model

The dual model begins with the specification of the individual's underlying indirect utility function. Dropping the individual subscripts for the moment let $H(p,y;Q,\theta,\varepsilon)$ denote the solution to a utility maximization problem defined by:

$$H\left(p,y;\mathrm{Q},\theta,\varepsilon\right)=\underset{z}{Max}\left\{U\left(z;Q,\theta,\varepsilon\right)\mid p'z+x=y\right\}, \tag{4}$$

where as above (z,x) denotes the private goods to be chosen by the individual, p is the vector of corresponding prices, y denotes income, and Q is a vector or matrix of public goods that the individual agent takes as given. In our application considering the demand for recreation the private goods include recreation trips to the available sites and all other spending, p reflects the costs of traveling to those sites, and Q includes the quality attributes of the sites. These attributes, while taken as given by the individual, include factors (such as congestion and fish stock) that are determined in equilibrium by the decisions made in the market as a whole. The direct utility function $U(z;q,\theta,\varepsilon)$ depends upon parameters θ and attributes of the individual ε that are unobserved by the analyst.

It is important to note that the indirect utility function in equation (4) is derived without imposing non-negativity constraints on demand. Applying Roy's Identity to equation (4) thus yields *notional* (or latent) demand equations

$$z_j^* = H_j\left(p,y;Q,\theta,\varepsilon\right) \Big/ \sum_{k=1}^{J} p_k H_k\left(p,y;Q,\theta,\varepsilon\right), \quad j=1,\dots,J, \tag{5}$$

where $H_j\left(p,y;Q,\theta,\varepsilon\right) \equiv \partial H\left(p,y;Q,\theta,\varepsilon\right)/\partial p_j$. Note that $z_j^* = z_j^*\left(p,y;Q,\theta,\varepsilon\right)$ will be negative for goods the individual does not wish to consume and positive for those goods she does consume. Observed consumption levels are then derived through the use of *virtual* prices, which rationalize the observed corner solutions. For example, suppose that the first r goods are not consumed. Let $p_N=(p_1,\dots,p_r)'$ denote the prices for the non-consumed (i.e., corner solution) goods and $p_C=(p_{1+r},\dots,p_J)'$ denote the prices for the consumed goods (i.e., those with positive consumption). The virtual prices for the non-consumed goods are implicitly defined by

$$0 = H_j\left[\pi_1\left(p_C;Q,\theta,\varepsilon\right),\dots,\pi_r\left(p_C;Q,\theta,\varepsilon\right),p_C,y;Q,\theta,\varepsilon\right], \quad j=1,\dots,r. \tag{6}$$

The observed demands for all the commodities become

$$z_j = H_j\left(p^*, y; Q, \theta, \varepsilon\right) \Big/ \sum_{k=1}^{J} p_k H_k\left(p^*, y; Q, \theta, \varepsilon\right) = z_j\left(p^*, y; Q, \theta, \varepsilon\right) \begin{cases} = 0 & j = 1, \ldots, r \\ > 0 & j = r+1, \ldots, M, \end{cases} \quad (7)$$

where $p^*=(\pi_1,\ldots,\pi_r, p_{1+r},\ldots,p_J)'$. The virtual prices are similarly linked to observed prices,

$$\pi_j\left(p^*, y; Q, \theta, \varepsilon\right) \begin{cases} \leq p_j & j = 1, \ldots, r \\ = p_j & j = r+1, \ldots, M. \end{cases} \quad (8)$$

The system of equations in (7) and (8) provides the connections between the observed data on usage (and prices) and the implied restrictions on the underlying error distributions, which can in turn be used for estimation. These equations allow us to express the indirect utility function (the solution to the utility maximization problem with non-negativity constraints enforced) as

$$V\left(p, y; Q, \theta, \varepsilon\right) = \max_{\omega \in \Omega} \left\{H(p^{\omega}, y; Q, \theta, \varepsilon)\right\}, \quad (9)$$

where Ω denotes the set of all possible *demand regimes* (combinations of corner and interior solutions among the J sites) and p^{ω} denotes the particular combination of virtual and actual prices associated with demand regime ω. The comparison of true and virtual prices as shown in equation (8) distinguishes the chosen goods from the non-chosen goods and thus can be used to gauge movements into and out of the market for particular commodities. In addition the virtual prices are functionally dependent on both the price and non-price attributes of the goods. Thus it is appropriate to view the virtual prices as quality-adjusted endogenous reservation prices, which can change in response to either price or non-price attribute changes.

B. *Model Specification*

The dual model specification can by obtained by choosing a functional form for the underlying indirect utility function $H(p,y;Q,\theta,\varepsilon)$ and deriving from it the corresponding notional demand and virtual price equations [17]. Alternatively, as in [34], one can begin with a system of notional demands for the goods of interest (i.e., an incomplete demand system) and integrate back to obtain the underlying quasi-indirect utility function. The advantage of the latter approach is that tractable notional demand equations can be specified, making the computation of virtual prices straightforward. The disadvantage here, of course, is that the resulting demand system is incomplete and does not capture substitution possibilities to goods outside of the choice set. In addition the restrictions needed to ensure integrability tend to require a choice between substitution and income effects.

Our empirical specification begins with the system of Marshallian notional demands

$$z_{ij}^{*} = \alpha_j + \sum_{k=1}^{J} \beta_{jk} p_{ik} + \gamma_j y_i + \varepsilon_{ij}, \quad j = 1,\ldots,J, \tag{10}$$

where the subscript i denotes people i=1,…,N and α_j=$\alpha_j(q_j,\xi_j)$ is a function of both observable site attributes for site j (the vector q_j – the component of Q specific to site j) and unobservable factors which we denote ξ_j. We assume that these factors have a linear form given by

$$\alpha_j = \gamma' q_j + \xi_j, \quad j = 1,\ldots,J. \tag{11}$$

The remaining notation is as follows: p_{ik} is the price of site k for individual i and y_i denotes individual i's income level. We impose a series of restrictions on the parameters of (10) to make the system of equations weakly integrable [16]. Specifically, we assume that $\beta_{jk} = \beta_{kj} \ \forall j,k$ and $\gamma_j = 0 \ \forall j$. The resulting notional demand system becomes:

$$z_{ij}^{*} = \alpha_j + \sum_{k=1}^{J} \beta_{jk} p_{ik} + \varepsilon_{ij}, \quad j = 1,\ldots,J. \tag{12}$$

The corresponding (notional) quasi-indirect utility function for (12) is given by:

$$\tilde{H}\left(p_i, y_i; Q, \theta, \varepsilon_i\right) = y_i - \left[\sum_{j=1}^{J}\left(\alpha_j + \varepsilon_{ij}\right) p_{ij} + \sum_{j=1}^{J}\sum_{k=1}^{J} \beta_{jk} p_{ij} p_{ik}\right]. \tag{13}$$

As indicated above the notional demand equations can be used to define virtual prices for the non-consumed goods that rationalize the observed corner solutions. For illustration, suppose again that the first r goods are not consumed. The virtual prices for these commodities are then

$$\pi_{iN} = -\beta_{NN}^{-1}\left[\alpha_N + \beta_{NC}p_{iC} + \varepsilon_{iN}\right] \tag{14}$$

where $\alpha_N=(\alpha_1,...,\alpha_r)'$, $\pi_{iN}=(\pi_{i1},...,\pi_{ir})'$, $p_{iC}=(p_{i,r+1},\ldots,p_{iJ})'$, $\varepsilon_{iN}=(\varepsilon_{i1},...,\varepsilon_{ir})'$,

$$\beta_{NN} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1r} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{r1} & \beta_{r2} & \cdots & \beta_{rr} \end{bmatrix} \text{ and } \beta_{NC} = \begin{bmatrix} \beta_{1,r+1} & \beta_{1,r+2} & \cdots & \beta_{1J} \\ \beta_{2,r+1} & \beta_{2,r+2} & \cdots & \beta_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{r,r+1} & \beta_{r,r+2} & \cdots & \beta_{rJ} \end{bmatrix}. \tag{15}$$

Note that the virtual prices π_{iN} depend on the quality attributes of the non-consumed goods. Substituting π_{iN} for p_{iN} in (12) yields observed demand equations for the consumed goods

$$z_{iC} = \alpha_C + \beta_{CN}\pi_{iN} + \beta_{CC}p_{iC} + \varepsilon_{iC} = \tilde{\alpha}_C + \tilde{\beta}_{CC}p_{iC} + \tilde{\varepsilon}_{iC}, \tag{16}$$

where $\tilde{\alpha}_C \equiv \alpha_C - \beta_{CN}\beta_{NN}^{-1}\alpha_N$, $\tilde{\beta}_{CC} = \beta_{CC} - \beta_{CN}\beta_{NN}^{-1}\beta_{NC}$, and $\tilde{\varepsilon}_{iC} = \varepsilon_{iC} - \beta_{CN}\beta_{NN}^{-1}\varepsilon_{iN}$. Notice that the observed demands depend directly only on the prices of those goods consumed, but that they depend upon the quality attributes for all the goods since all the α_j's enter into equation (16) and they each depend in turn upon the corresponding quality attributes.

Finally, in our empirical analysis below, we impose two additional simplifying restrictions on the model in equation (12). First, we assume that all of the own and cross-price coefficients are the same across sites; i.e., $\beta_{jj} = \beta_1 \ \forall j$ and $\beta_{jk} = \beta_2 \forall j \neq k$. Second, we assume that the idiosyncratic individual heterogeneity captured by $\varepsilon_i=(\varepsilon_{i1},...,\varepsilon_{iJ})'$ is *iid* $N(0,\Sigma)$, where $\Sigma = diag\left(\sigma_1^2,\ldots,\sigma_J^2\right)$. Neither of these restrictions is necessary from a conceptual perspective, but they substantially reduce the number of parameters that must be estimated.

V. Application

Our empirical analysis focuses on modeling the demand for lake recreation in Iowa, drawing on data from the first year of the Iowa Lakes Project (see [2] for an overview). This project is a four-year panel data study, analyzing the visitation patterns to 129 Iowa lakes by 8000 randomly selected households, providing a rich source of variation in usage patterns. Iowa is particularly well suited for our research for several reasons. First, Iowa's lakes are characterized by a wide range of water quality, including both some of the cleanest and some the dirtiest lakes in the world. Second, detailed information is available on the environmental conditions in each lake, including physical and chemical measures (e.g., Secchi transparency, total nitrogen, etc.) obtained three times each year during the course of the project. Third, Iowa is currently considering major lake water remediation efforts. These include multi-million dollar projects to improve target lakes as well as the Governor's stated objective to remove all of Iowa's lakes from the US EPA's impaired water quality list by 2010. These changes provide natural, and policy relevant, sets of scenarios to consider in investigating the general equilibrium effects of regulatory interventions.

The 2002 Iowa Lakes Project survey was administered by mail to a randomly selected sample of 8000 households. A total of 4423 surveys were completed, for a 62% response rate once non-deliverables surveys are accounted for. The key section of the survey for the current analysis obtained information regarding the respondents' visits to each of 129 lakes during 2002. On average, approximately 63% of Iowa households were found to visit at least one of these lakes in 2002, with the average number of day trips per household per year being 8.1. There are, of course, a large number of corner solutions in this dataset, with 37% of the households visiting none of the lake sites and most households choosing only a small subset of the available sites. Fewer than 10% of those surveyed visited more than five distinct sites during the course of a year. For the purposes of the econometric analysis below, a sub-sample of the 2002 usage data was used. Specifically, we had available 1286 observations, randomly selected from the full sample. These records were further narrowed to the 749 users in the sub-sample (i.e., households taking at least one trip in 2002).

The Iowa Lakes survey data was supplemented with two data sources. First, for each individual in the sample, travel distance and time from their home to each of 129 lakes were calculated using the transportation software package PCMiler. Travel costs were then computed using a distance cost of \$0.28/mile and valuing the travel time at one-third the individual's wage rate. The average travel cost over all site/individual combinations was \$135. Second, water quality measures for each of the lakes were provided by the Iowa State Limnological Laboratory. The average Secchi transparency and Chlorophyll levels are used in the current analysis. Secchi transparency measures water clarity and ranges from 0.09 to 5.67 meters across the 129 lakes in our sample. Chlorophyll is an indicator of phytoplankton plant biomass which leads to greenness in the water, and ranges from 2 to 183 μg/liter among the Iowa lakes.

VI. Estimation Algorithm

The estimation of the parameters in our model of recreation demand can be divided into two stages. In the first stage, we estimate the basic parameters in the demand system characterized by equations (7) and (8). Specifically, we obtain posterior distributions for the $2J+2$ parameters $\theta=(\alpha_1,...,\alpha_J, \beta_1, \beta_2, \sigma_1,...,\sigma_J)$ using a Bayesian computational approach relying on Gibbs sampling and data augmentation. Note that site specific intercept terms capture all of the site specific attributes, including the endogenous factors of interest and unobserved site characteristics. The purpose of the second stage then is to estimate the functional relationship between these intercepts and the observed quality attributes.

A. First Stage Estimation

Data augmentation techniques pioneered by [1] in the context of discrete choice models provide a powerful tool for handling latent variables, simulating these missing components of the data and, in doing so, making the analysis more tractable. In the current context, the latent variables are the notional demands. Together with Gibbs sampling techniques we can use data augmentation to simulate values from the posterior distribution of interest.[2]

Formally, the posterior distribution is characterized by

$$p\left(z^{*},\theta \mid z\right) \propto p\left(z \mid z^{*},\theta\right) p\left(z^{*} \mid \theta\right) p\left(\theta\right). \tag{17}$$

Note that the data augmentation procedure treats the unknown latent factors essentially as additional parameters, characterized in terms of a prior distribution and for which a posterior distribution is generated. The priors for θ are assumed to take the following forms:

$$\psi \equiv \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \sim N\left(\bar{\psi},\tau I_m\right), \tag{18}$$

where τ is a large constant and I_m is an $m \times m$ identity matrix with $m=J+2$, and the σ_j^2 are independent inverted gamma variates with $\sigma_j^2 \sim IG\left(1,\underline{s}_j\right)$. While the joint posterior distribution in (17) is complex, the corresponding conditional posterior distributions for z^*, ψ, and σ_j^2 each have convenient functional forms. Details of the iterative Gibbs sampling routine used to simulate draws from the posterior distribution are provided in Appendix A to this paper.

B. Second Stage Estimation

The first stage provides realizations from the posterior distribution for the parameters $\alpha_1,\ldots,\alpha_J$ as well as the price and error variance terms. In this subsection we are interested in decomposing the intercepts into components that represent the observable and unobservable attributes of the recreation sites. Since the former will include endogenous attributes that ultimately drive the general equilibrium aspects of the model this decomposition is important. Likewise any welfare analysis concerning site attributes (partial or general equilibrium) will require understanding how demand is impacted by changes in observable site quality levels.

The linearity of equation (11) suggests that the unknown parameter vector γ can be computed via a regression of the intercepts on the observed quality attributes q_j, with ξ_j then computed as the residual from the regression estimates (see [20,30]). We nest this notion within the Bayesian estimation paradigm as follows. For each of the 800 draws from the posterior distribution of $\alpha_1,\ldots,\alpha_J$ obtained in the first stage we regress the realized intercept values on the vectors of site attributes $q_1,\ldots,q_J$. For each draw of the intercepts we therefore obtain a value for γ. The set of these values gives an empirical distribution that characterizes the posterior distribution for the unknown parameters γ, which provides measurements of how observable site attributes affect site demands.

If an endogenous attribute is included in the vector of attributes q_j then steps must be taken to account for the likely econometric endogeneity in the second stage regression. For the case of congestion, a site is likely to be heavily visited (and hence congested) if it possesses attractive attributes, some of which will not be measured by the econometrician and will therefore reside in ξ_j. This will induce (positive) correlation between the measure of congestion and the error in the second stage regression, leading to biased estimates for the role of all observable attributes on site demand. We deal with the endogenous congestion attribute via a strategy recently suggested by [30] in the context of a site choice model. We define our baseline measure of congestion as the share of people who visit a particular site. Thus the baseline share of people who visit site j is defined by $s_j^0 = N^{-1}\sum_{i=1}^{N} I_{ij}$, where I_{ij}=1 if the person i's notional demand for site j is positive (i.e. $z_{ij}^* > 0$) and zero otherwise (i.e. $z_{ij}^* \leq 0$). Our second stage estimation problem for each posterior draw from the intercepts is therefore given by

$$\alpha_j = \gamma_0 + \sum_c \gamma_c q_{jc} + \gamma_s s_j^0 + \xi_j, \quad j = 1,...,J, \tag{19}$$

and an instrumental variable is needed for s_j^0. Appendix A provides additional details on how we construct an instrument and complete the second stage analysis.

C. Welfare Analysis

In this section we describe how welfare analysis of changes in prices or attributes at the recreation sites proceeds. As discussed in [23,31,32], welfare analysis in this class of models is complicated by many factors, including the need to simulate behavior under counterfactual conditions. The steps needed for analysis using the dual model are particularly challenging and involve subtle technicalities and many computational issues that at this stage of research are not fully understood. The added complexity of computing general equilibrium welfare measures compounds the difficulties. Thus we focus in this section on laying out the broad steps needed for computing welfare measures, forgoing many of the technical details in favor of a discussion of the conceptual challenges that we have solved and those that must still be examined.

The first and second stages of the estimation algorithm provide summaries via posterior means of the utility function parameters and the error variances that constitute the unknowns in the model. From these summaries preferences are characterized up to the values of the idiosyncratic errors (i.e., the ε_{ij}'s). Because the idiosyncratic errors are given a structural interpretation the utility function is random from analyst's perspective, which implies that the compensating surplus for changes in prices or site attributes will also be random variables. Thus the objective of welfare analysis is to estimate the expected value of an individual's compensating surplus by integrating out the unobserved determinants of choice. This requires that we simulate values of the idiosyncratic errors for each person many times and compute the welfare measure of interest for each simulated draw. Averaging over the welfare measures for each simulated error draw provides an estimate of each person's expected welfare measure.

Conditional on a simulated value of the errors for each person all components of the preference function $V(p,y;q,\theta,\varepsilon)$ shown in (9) are known. In general we can define compensating surplus (CS) implicitly using $V(\cdot)$ by

$$V(p^0, y, Q^0, \theta, \varepsilon) = V(p^1, y - CS, Q^1, \theta, \varepsilon), \tag{20}$$

where the change is being evaluated from baseline conditions (p^0,Q^0) to new conditions (p^1,Q^1). Solving for CS in equation (20) is complicated by the fact that the indirect utility function represents an endogenous regime-switching outcome: changes in p or Q may induce visitors to change the pattern of sites they visit as well as the frequency, and the solution algorithm must ensure that the comparison is made for utility values that reflect the appropriate (utility maximizing) demand regime under initial and changed conditions.

Simulating the errors and solving the consumer's problem are required for both partial and general equilibrium analysis and involve similar computation steps and techniques. The general equilibrium calculation adds an additional layer of computation in that the simulated behavior at new conditions must also be used to predict the elements of Q^1 that are endogenously determined by the aggregation of behavior. Thus welfare calculation in this paper requires we address two methodological challenges: solving the consumer's problem when preferences are not additively separable and predicting new levels of the endogenous attributes under changed conditions. We must also address the non-linear way in which site quality attributes enter V (recall they enter indirectly through the virtual prices for the non-consumed goods as well as directly via the intercepts) to properly compute and interpret our welfare measures.

To illustrate these challenges and describe our initial solutions we first list out the explicit steps needed to compute partial and general equilibrium welfare measures using the specification we are working with. We then describe the steps in more detail. Given the posterior means for the utility and error distribution parameters partial equilibrium welfare analysis for a single person i and a single draw of the errors involves the following steps:

1) Draw values of the errors $\varepsilon_{i1},\ldots,\varepsilon_{iJ}$ from the estimated distribution for the unobserved component of utility *conditionally* such that observed levels of demand at baseline conditions are replicated for person i. That is, draw errors such that

$$z_{iC}^0 = \alpha_C^0 + \beta_{NC}\pi_{iN}^0 + \beta_{CC}p_C^0 + \varepsilon_{iC}, \tag{21}$$

where the notation follows from equation (16), superscripts '0' indicate that all price and quality variables are set to their baseline values, and z_{iC}^0 denotes the observed level of visits for person i to the set of visited sites C.

2) Determine the total baseline consumer surplus *from the consumed goods* by integrating under the baseline demands in equation (21) between p_{iC}^0 and $\hat{p}_{iC}^0$, where $\hat{p}_{iC}^0$ is the choke price for the set of goods C at baseline conditions. Because the ordinary and compensating demand curves are the same in our case this is also the total Hicksian consumer surplus.

3) Define the counterfactual scenario to consist of new prices and/or exogenous site attribute levels $p_{i1}^1,\ldots,p_{iJ}^1$ and $\alpha_1^1,\ldots,\alpha_J^1$, where

$$\alpha_j^1 = \gamma_0 + \sum_c \gamma_c q^1{}_{jc} + \gamma_s s_j^0 + \xi_j, \quad j = 1,\ldots,J, \tag{22}$$

and the q_{jc}^1's hold new levels of the exogenous attributes. For a partial equilibrium analysis the original levels of the endogenous attributes (congestion) are maintained.

4) Determine the new demand regime by computing the new notional demands using

$$z_{ij}^{*1} = \alpha_j^1 + \sum_{k=1}^{J} \beta_{jk} p_{ik}^1 + \varepsilon_{ij}, \quad 1,...,J, \tag{23}$$

and observe the pattern of positive and negative values for z_{ij}^{*1}. The combination of positive and negative values for each site is a candidate $\tilde{C}^1$ for the new demand regime, which must be evaluated and updated as described below. Via the updating the new demand regime C^1 is determined.

5) Determine the total baseline consumer surplus from the consumed goods at the *new demand regime* by integrating under $z_{iC^1}^1 = \alpha_{C^1}^1 + \beta_{NC^1}\pi_{iN^1}^1 + \beta_{C^1C^1} p_{C^1}^1 + \varepsilon_{iC^1}$ between $p_{iC^1}^1$ and $\hat{p}_{iC^1}^1$, where $\hat{p}_{iC^1}^1$ is the choke price for the set of goods C^1 at changed conditions.
6) Compensating surplus for person *i* for this draw of the error is the difference between total surplus at initial and changed conditions.

A few comments on this algorithm should be made. First, the algorithm focuses on obtaining use-only values from changes in the levels of attributes by only including areas under the demand curves for sites that are actually visited. A utility function approach as shown in general in equation (20) would also add surplus to the total for sites that were not visited. We have chosen to focus on use value to aid in interpretation and avoid complexities associated with general equilibrium feedbacks interacting with non-use value computation. Also, we are assuming in this algorithm that changes in observable attributes from Q^0 to Q^1 leave unobserved attribute levels unchanged. That is, ξ_j is constant for all sites across all changes. Depending on the scenario this may or may not be a realistic assumption.

For general equilibrium welfare measurement the same steps are needed, but we must also update the level of the endogenous attribute. Step 3' for the GE case becomes

3’)Define the counterfactual scenario to consist of new prices and *candidate* site attribute levels defined by $p_{i1}^1,...,p_{iJ}^1$ and $\tilde{\alpha}_1^1,...,\tilde{\alpha}_J^1$, where

$$\tilde{\alpha}_j^1 = \gamma_0 + \sum_c \gamma_c q^1{}_{jc} + \gamma_s s_j^0 + \xi_j, \quad j = 1,...,J, \tag{24}$$

and the q_{jc}^1's hold new levels of the exogenous attributes. For general equilibrium measurement an algorithm is needed that updates $\tilde{\alpha}_1^1,...,\tilde{\alpha}_J^1$, to $\alpha_1^1,...,\alpha_J^1$, where

$$\alpha_j^1 = \gamma_0 + \sum_c \gamma_c q^1{}_{jc} + \gamma_s s_j^1 + \xi_j, \quad j = 1,...,J, \tag{25}$$

and s_j^1 is the new equilibrium proportion of people who visit site *j* given the changed conditions. We discuss the form that this updating takes below.

These six steps each present varying degrees of computational and conceptual challenges. Step 1 is technical but quite similar to the data augmentation stage described above. In addition the ideas associated with simulating unobserved errors consistent with observed choice are well-explained in other work. Steps 2 and 5 are likewise mechanical and we forgo additional discussion of these steps. We do note however that the decision to rely on surplus measures computed as areas under compensated demand curves rather using expenditure or utility functions is a point worthy of further discussion, but one that is largely orthogonal to the topics directly under consideration. Thus the steps that we provide more discussion on include step 4 (determining the new demand regime) and step 3 as it relates to the general equilibrium calculation.

Consider first how we determine the new demand regime given changed conditions. We refer to equation (23) as providing a 'candidate' demand regime because the mapping between a set of positive and negative *notional* demands to the implied *actual* demands via the appropriate set of virtual prices does not guarantee that the resulting actual demands will be strictly positive. Thus the candidate regime may not be a member of the set of feasible demand regimes under the changed conditions. In this case a mechanism is needed to find an alternative demand regime from the set of feasible regimes (i.e. those that do not result in negative actual demands) that maximizes utility. We have not yet solved this problem formally and rely at this stage on an ad hoc updating rule. Specifically, we complete step 4 via the following:

- Observe the candidate regime $\tilde{C}^1$ using equation (23).
- Compute the candidate actual demands $\tilde{z}^1_{i\tilde{C}^1}$ for this regime and observe which, if any, of the demands are negative.
- Update the candidate demand regime by setting to 'non-consumed' the goods observed with negative actual demands. Label this C^1 (the new regime) and use it in step 5.

This is an ad hoc updating rule in that we have not proven that it results in the utility maximizing solution under the new price and attribute conditions. Verifying this to be the case, or altering the updating strategy to find the maximizing solution, is an important area for subsequent research.

Step 3 is trivial in the partial equilibrium case but involves notable challenges for the general equilibrium case. An updating rule is needed that re-equilibrates the site quality indexes (the intercepts) to reflect both the new exogenous attribute levels and the new resulting congestion level. A similar challenge was faced by Timmins and Murdock [30] for their site choice congestion application. These authors rely on results in [6] to show that their measure of congestion is the result of a unique sorting equilibrium, and that solving for new congestion levels in counterfactual experiments relies on a simple application of Brower's fixed point theorem. Since our measure of congestion and behavioral model differ from [30] these results do not transfer directly. Thus at this stage of the research we are still investigating the formal properties of our equilibrium as well as computational methods for simulating new outcomes.

To explore this point further recall that our measure of congestion as given by s_j^0 consists of the proportion of people who visit a particular site. This is a convenient metric for our model in that the related concepts of virtual price and notional demands can in principle be used to predict this proportion for the sample under any configuration of prices and exogenous attributes. Consider the following algorithm. For iterations t=1,2,… complete the following steps:

a) Define $\tilde{z}_{ij}^{*t}$ by

$$\tilde{z}_{ij}^{*t} = \tilde{\alpha}_j^{1t}(\tilde{s}_j^{t-1}) + \sum_{k=1}^{J} \beta_{jk} p_{ik}^1 + \varepsilon_{ij}, \quad j = 1,\dots,J, \quad i = 1,\dots,N,$$

$$\tilde{s}_j^{t-1} = \frac{1}{N}\sum_{i=1}^{N} I_{ij}^{t-1}, \quad I_{ij}^{t-1} = \begin{cases} 1 & \tilde{z}_{ij}^{*t-1} > 0 \\ 0 & \tilde{z}_{ij}^{*t-1} < 0 \end{cases}$$

b) At iteration T define the new equilibrium congestion level by $s_j^1 = \frac{1}{N}\sum_{i=1}^{N} I_{ij}^{T-1}$.

Understanding the formal properties of this mapping, and altering it as needed, is an important task for further research.[3]

VII. Empirical Results

In this section we present empirical results for the Iowa lakes data set described above. We emphasize that at this stage these findings are illustrative and exploratory. Nonetheless several interesting results emerge that illustrate the importance of non-price equilibria concepts. The model was run in MATLAB using the first stage Gibbs sampler and second stage regression decomposition to obtain an empirical representation of the posterior distribution for the unknowns in the model. The first stage is computationally intense: obtaining 7200 draws from the posterior distribution required nearly a month of run time on a new computer. From the 7200 draws obtained we discard the first 4000 as a burn-in period and construct our empirical distribution using every fourth draw thereafter, leaving 800 draws of the 258 first stage parameters for inference and subsequent analysis.

Table 1 contains a summary of the posterior distribution for the own- and cross-price parameters.[4] The price coefficients can be directly interpreted as the marginal effects of price changes on the notional demands, but their interpretation as related to the actual demands is more complex since the parameters enter the demand equations non-linearly. Nonetheless the signs and ratios of the posterior means and standard deviations seem reasonable. We find that own price effects are two orders of magnitude larger than the cross price effects, suggesting there may be little cross-site substitution on average. Our parsimonious specification may, however, mask larger cross price effects between individual sites.

Tables 2 and 3 present different strategies and results for decomposing the intercepts into observable and unobservable site attributes. Once the empirical distribution for the intercepts is obtained in the first stage it is computationally fast and straightforward to investigate different specifications for the second stage. We experimented with different water quality attributes as second stage explanatory variables and settled on the use of two: Secchi disk measurement and ambient levels of chlorophyll. These two measures of site quality are potentially attractive in that their effects are observable to visitors. Secchi readings reflect observable water clarity (assumed to be a positive attribute of lakes) while chlorophyll reflect visible algae and weed growth, which are correlated with nitrification. We stress nonetheless that other specifications may be preferred. We are however degrees-of-freedom limited. The second stage regressions exploit variation over sites, so estimates are based in our case on only 128 observations.

The results are illustrative of both the difficulties and importance of accounting for endogenous attributes such as congestion. Table 2 contains our straw-man results. Here we have naively included an obviously endogenous variable (proportion of people visiting each site) in the equation and used OLS to estimate the parameters. We find a negative effect on congestion but have no resolution on our site quality estimates. In contrast the results in table 3 are much more promising and intuitive. We find a large and negative coefficient on congestion and a solidly significant coefficient of the correct sign on chlorophyll. The sign on Secchi is appropriately positive but at best marginally significant. We cautiously conclude that our instrument strategy is viable and that congestion matters – probably more than exogenous attributes such as water quality and perhaps as much as own price effects. This finding is similar to [30], who find using their preferred instrument strategy large and significant disutility from congestion.[5]

A primary objective of estimating the parameters of the structural model is to examine both partial and general equilibrium welfare measures. To illustrate the capabilities of the model in this dimension we again consider four counterfactual scenarios, each designed to illustrate welfare measures of potentially different types. The scenarios are:

Scenario 1: Close the most heavily visited lake in each of nine regions of Iowa.

Scenario 2: Close a moderately visited site in each of nine regions of Iowa.

Scenario 3: Improve water quality throughout the state such that all lakes obtain at least the rating of 'good water quality'. This corresponds to a minimum Secchi reading of 2.17 meters and maximum chlorophyll reading of 8.26ug/l.

Scenario 4: Improve a set of seven Iowa Department of Natural Resources 'target lakes' to a minimum Secchi reading of 5.7 meters and maximum chlorophyll reading of 2.6ug/l.

The first scenario is major in that it involves the loss of nine primary lakes in the state, while the second is arguably minor in that the lakes are minor regional facilities. In both cases we proxy the loss of the sites by setting travel costs above the choke prices for all visitors in the sample. The third and fourth scenarios are minor but widespread and major but localized, respectively.

Point estimates for both partial and general equilibrium compensating surplus measures are shown in Table 4 The estimates are seasonal per person measures; for a rough per trip measure one could normalize by the mean or median total annual trips taken (11.46 and 7 trips, respectively). We find plausible estimates for the partial equilibrium estimates in all cases. The loss of the nine popular sites leads to a large mean welfare loss of over $600 per person per season. The much smaller median loss of nearly $68 per person suggests the mean is skewed by individuals with high valuations for the lost sites. This is sensible in this model and matches our intuition: people who do not visit the lost sites, or do so only infrequently, do not suffer a surplus loss when the sites are eliminated. This is also seen in scenario 2, where the sample average loss from closing 9 moderately popular sites is nearly $52 per angler and the median is zero. Closing the less important sites impacts less than half the people in the sample, suggesting that most suffer no surplus loss from the closures.

Scenarios 3 and 4 examine quality improvements of different intensity and spatial extent. Here we find that smaller, more widespread quality improvements have a larger welfare impact (sample mean and median of approximately $208 and $147, respectively) than their localized but larger counter parts ($50 and $17). In both cases the positive median suggests benefits are spread throughout the sampled population.

The general equilibrium welfare measures also seem plausible for all scenarios, though we caution that these measures are preliminary in that further research is needed to understand the properties of re-equilibrating algorithm. For the site loss scenarios in particular it is unclear that a new sorting equilibrium is achieved; evidence of this is stronger for the quality changes. Nonetheless some intuition emerges. Using the means in scenario 1 we find general equilibrium welfare losses that are 15% larger than their partial equilibrium counterpart. Similarly for scenario 2 we find general equilibrium losses that are 11% higher. This reflects the fact that the site closures cause a direct welfare effect via the lost choice alternatives as well as an indirect effect on the remaining sites via increased congestion. The intuition for the direction of the general equilibrium effect is less clear for the quality changes. If improvements are made to moderately popular lakes, and by attracting visitors from more popular lakes decreases congestion, the general equilibrium effect may be larger. In contrast improvements at currently congested sites that cause more people to go to these sites may lead to smaller general equilibrium welfare improvements. For our scenarios we find general equilibrium effects that suggest smaller improvements when re-sorting is accounted for.

VIII. CGE Modeling

A virtue of our empirical model is its richness. It is designed to take full advantage of the micro-level information that is available in datasets such as ours. Therefore, if one accepts the underlying structure of the model it is the natural choice for obtaining parameter estimates and forecasting the equilibrium response to policy changes. However, a number of assumptions in the model are speculative. CGE modeling provides a flexible way to explore the behavior of non-price equilibrium models under alternative specifications and a robust set of tools for solving equilibrium problems with non-market feedback effects. Carbone and Smith [10] demonstrate this in their analysis of air pollution externalities. Here we develop a CGE model for the Iowa Lakes application that complements the econometric analysis. The objective is to demonstrate how CGE can be used in this setting and to test the consequences of altering two of the underlying assumptions in the empirical model.

Our attention in the counterfactual analysis of the empirical model focused on the differences between welfare estimates produced by either a partial or general equilibrium version of the model, where these differences were driven by the magnitude of the feedback effects between site demand and congestion levels. Within the structure of our model the two most obvious drivers of this mechanism are the magnitude of the congestion demand coefficient, and the form of the congestion transmission function. These assumptions are the focus of our sensitivity analysis. By design, the CGE model shares many common features with the empirical model, so we forego a discussion of these aspects of the problem in favor of one that highlights the differences in model structure, solution concept, and calibration.

A. Model Structure and Solution Concept

The set of sites and quality characteristics remain unchanged in the CGE model. However, due to computation time constraints we do not represent the full population of individuals. Instead we present the results of models based on a sub-sample of 300 randomly selected individuals from the data. The results presented below are based on one such sub-sample. Following the logic of the estimation strategy, the demand system is based on a dual representation. The outcome of the consumer choice process can be represented by a conditional indirect utility function

$$\tilde{H}\left(\pi_i, y\right) = y - \left[\sum_j \left(\gamma_{ij}^0 + \sum_m \gamma_m q_{mj}\right)\pi_{im} + \tfrac{1}{2}\sum_j \sum_k \beta_{jk}\pi_{ij}\pi_{ik}\right], \qquad (26)$$

where π_i is a vector of virtual prices that rationalizes the regime-specific system of notional demands to describe the actual demand system. Roy's Identity allows us to derive the actual demand equations and their relationship to the virtual prices:

$$\frac{-\partial \tilde{H}_i / \partial \pi_{ij}}{\partial \tilde{H}_i / \partial y_i} = z_{ij} = \gamma_{ij}^0 + \sum_m \gamma_m q_{mj} + \sum_k \beta_{jk}\pi_{ik} \geq 0 \;\perp\; \pi_{ij} \leq p_{ij}, \qquad (27)$$

where $\perp$ indicates the complementary slackness relationship between the non-negativity constraint on site demand on the left-hand side of (27) and the constraint on the virtual price on the right-hand side. If the non-negativity constraint on demand for site j is binding, then the constraint on the virtual price may be slack. If the non-negativity constraint on demand for site j

is slack, however, the virtual price constraint must bind. This relationship is at the heart of the solution concept for the CGE model. Congestion makes its contribution to utility separately from the other quality characteristics in the data and these characteristics are assumed to be exogenous such that $q_{mj} = q_{mj}^0 \;\forall m \neq 1,$ m=1 denotes our congestion measure, and q_{mj}^0 is the baseline level of the attribute observed in the data for each site in the model.

The solution methods employed in both the empirical model in this paper and previous experiments based on the random utility model [30] require that we define the measure of site congestion as the share of visitors who visit a given site. The CGE model allows us to relax this assumption and explore specifications for which the intensity of visitation affects congestion levels even in the absence of re-sorting across sites. We consider two different specifications for the transmission function. In the first, congestion is defined as the aggregate demand share at site j:

$$q_{1j} = \sum_i z_{ij} \left[\sum_l \sum_k z_{lk} \right]^{-1}. \qquad (28)$$

This specification mimics the approach used in our empirical model and in [30]. The second specification defines congestion as simply the total number of visitors to site j:

$$q_{1j} = \sum_i z_{ij} \qquad (29)$$

This specification takes the intensity of visitation on congestion into account. In the discussion that follows, we refer to versions of the model that use (28) as the “shares” model and the versions that use (29) as the “totals” model.

The full equilibrium model consists of first-order conditions based on (27) and the congestion transmission functions described in either (28) or (29). This system of equations and complementary slackness relationships belongs to a class of problems referred to as the Mixed

Complementarity Problem (MCP). The GAMS software [9] combined with the PATH solver [13] provides a robust way to solve large-scale MCPs.

B. Equilibrium Effects of Congestion

Having specified functional forms, we can provide intuition for how the congestion feedback effects functions and motivation for the second part of the sensitivity analysis which involves altering the value of the congestion demand coefficient. Re-writing (29) we have

$$q_{1j} = \sum_i \left(\gamma_{ij}^0 + \sum_m \gamma_m q_{mj} + \sum_k \beta_{jk} \pi_{ik} \right). \tag{30}$$

Collecting the q_{1j} terms on the left-hand side of the equation and solving for q_{1j}, we have

$$q_{1j} = \sum_i \left(\gamma_{ij}^0 + \sum_{m \neq 1} \gamma_m q_{mj} + \sum_k \beta_{jk} \pi_{ik} \right) \Big/ \left(1 - I\gamma_1\right). \tag{31}$$

The denominator is greater than one if congestion has a negative effect ($\gamma_1 = \gamma_{congest} < 0$) on visitation. Now consider the effect of a policy change on the equilibrium level of congestion. If, for example, some aspect of water quality is improved at site j, the direct effect on the level of congestion, measured by the numerator of (31), causes the level of congestion at site j to rise. However, the negative feedback effect caused by this increase in congestion tends to decrease demand for site j which, in turn, reduces the level of congestion. This process of adjustment continues until we reach a fixed point. The expression in (31) captures the net effect of this adjustment process on the level of congestion, where the denominator term describes the factor by which the direct effect of the increase in congestion is attenuated by the general equilibrium feedback effects.[6] Notice that the size of the feedback effect depends directly on the magnitude of the $\gamma_{congest}$ parameter. We explore this link in the sensitivity analysis that follows.

C. Calibration and Counterfactual Experiments

The objective of our calibration is to match the CGE model to the observed demands in the Iowa Lakes data at the baseline equilibrium, and to use the point estimates of the demand coefficients in the estimation routine to calibrate the price and quality responsiveness of the individual demand functions. The point estimates that are used are: $\beta_{jj}=-0.045$ for all j, $\beta_{kj}=0.0002$ for all $j \neq k$, $\gamma_{secchi}=0.084$, $\gamma_{chloro}=-0.006$, and $\gamma_{congest}=-54.34$. When we use the totals specification, $\gamma_{congest}$ is re-scaled by dividing by the total number of visits to all sites in the benchmark dataset.

A premise of the estimation strategy is that unobserved components of individual tastes drive site demands and lead to corner solutions for particular sites. Conceptually, the calibration of tastes in the CGE model follows the same logic used to simulate tastes in the empirical model and we refer the reader to step 1 of the welfare analysis procedure described in section VI.C for a description of this process. The use of this technique in a CGE model is, to our knowledge, new and potentially of interest in its own right. Details on this and other aspects of the calibration routine are provided in Appendix C.

As in the empirical model, the discussion in the sensitivity analysis focuses on the difference between partial equilibrium (PE) measures of change in consumer surplus in which the level of congestion at each site is held constant at its benchmark level and general equilibrium (GE) measures in which site demand and congestion levels are both endogenously determined in equilibrium. The methods for the welfare calculations follow the logic described in section VI.C with one exception. Determining the counterfactual demand regimes (step 4) is accomplished by passing the MCP description of the equilibrium model and parameter values that are consistent with the counterfactual policy to PATH, the MCP solver in GAMS.

The welfare results that we report are based on the same counterfactual scenarios used in the empirical model and described in section VII. For each of the policy scenarios that we have outlined, we simulate the calibrated model under the assumption that the $\gamma_{congest}$ parameter takes on values of one half, one, and two times the value of the point estimate for this parameter from the estimation results. The benchmark model uses the totals specification of the congestion transmission function. We compare welfare estimates from this model with the results of the same policy experiments performed in the shares model.

D. Results

Table 5 reports on the welfare results for the core simulations. The rows list the four different policy scenarios. The columns of the table indicate the welfare measures, PE or GE, and the percentage difference between the two. The columns also indicate the assumed value of $\gamma_{congest}$ .The GE welfare measures, which take into account the adjustments in site congestion levels, differ by only 2% from the PE estimates in magnitude for the policy scenarios that involve shutting down sites (scenarios #1 and #2), suggesting that the congestion response is small. However, the differences are more substantial for both of the water quality improvement scenarios (#3 and #4) – roughly 40% in both cases. Changing the value of $\gamma_{congest}$ parameter has the hypothesized effect. The differences between the PE and GE welfare measures are magnified when the value of this parameter is large.

The results of the quality change experiments are relatively intuitive. The direct effect of the policy is to increase the quality of targeted sites. This leads to welfare gains to visitors. The general equilibrium response is higher congestion at improved sites which tends to offset some of the PE welfare gains. In light of these results, it is somewhat surprising that the feedback effects are so much smaller in the site-shutdown scenarios. Table 6, which describes the results of the comparison between the shares and the totals models, is helpful in explaining this outcome. As before, the rows of the table describe the different policies. The columns describe

PE and GE welfare measures for each of the two different versions of the model. In the shares model, the policies that were left unchanged by congestion effects under the totals assumption (site-shutdown scenarios) lead to welfare losses in general equilibrium that are roughly 6% larger than the PE welfare estimates. More striking, however, is the result that the large impacts of congestion effects in the quality-improvement scenarios that we see in the totals model disappear in the shares results.

Site shutdown is equivalent to an increase in the price of targeted sites. This price increase implies substitution into alternative sites. This leads to higher congestion levels at these sites. Thus, it is substitution across sites that drives the differences between the PE and GE welfare measures in these experiments. The cross-price coefficient from the econometric results (β_{kj}=0.0002) is relatively small. Therefore, we would expect the congestion effects driven by site substitution to also be small. In contrast, the quality-improvement scenarios have more scope for congestion effects generated by the increased intensity of visitation rather than substitution. The coefficient on water quality (γ_{secchi}=0.084) is relatively large, so the demand response and congestion response to these types of policies can be quite large as well. Notice that this latter type of congestion effect – driven by visit intensity – is only possible in the totals specification of the model, where the congestion transmission function captures intensive margin adjustments. In the shares version of the model where intensity does not affect congestion, the PE-GE differences for these policies disappear.

Overall, both dimensions of the sensitivity analysis yield insights. In particular, the comparison of the results from the totals and shares models and the interaction of these models with the different policy scenarios demonstrate the wide variety of behavior that can be produced in even a fairly simple non-price equilibrium model such as the one we have described here.

More generally, our findings lead us to conclude that obtaining the form for these transmission functions in different policy applications should be a priority for future research in this area.

IX. Discussion

Our objective has been to explore the notion of non-price equilibria and feedback effects in non-market valuation using both CGE and econometric methods. We have focused on defining in general the concept of non-price equilibra, investigating the circumstances under which they might be empirically important for non-market valuation, and exploring how to measure the effects if they exist. As we stressed above this paper is the initial rather than the final step in this direction, and it leaves much unresolved. Nonetheless several insights have emerged along with promising leads for continuing this line of research. To conclude the paper we summarize the findings, lessons, and speculations that have emerged and identify specific areas for further research, placing them in the context of the 'frontiers' theme of this paper.

While we have categorized the types of non-price equilibria that may arise using the concepts of simple and complex sorting equilibrium it is difficult to say much more of operational value that is not context specific. Nonetheless there is a fairly general set of issues that have emerged and need to be resolved in any application of this type. These include, for example, specifying the mechanism through which aggregate behavior translates into an endogenous attribute, choosing a tractable parameterization for this mechanism (i.e. the transmission function), and determining the computational and conceptual properties of the equilibrium associated with the transmission function. We have provided examples of these issues as they relate to congestion in a fairly general model of seasonal multiple site recreation demand.

To investigate congestion in the recreation context we have taken the fairly unusual step of using both CGE and econometric modeling approaches. Our intent, and that of the workshop organizers who suggested and helped assemble the research team, was to consider the issue from different perspectives using what we hoped would be complementary tools. This proved to be a good decision and gives an example of the scientific value of coordinated experiments and replication exercises.

The CGE model functioned as a laboratory which, when calibrated to our application, allowed us to explore how partial and general equilibrium results are sensitive to the magnitude of the congestion effect and the parameterization of the transmission function. Both of these dimensions of the problem turned out to be important. The flexibility of this approach, both in decomposing the results of counterfactual scenarios and performing sensitivity analyses, makes it a valuable tool when the modeling exercise demands that we represent multiple, interacting channels of influence. The role for this type of analysis becomes even clearer when one considers that we have chosen a relatively simple example of a non-price equilibrium as our application. A necessary precursor to any experiment in which we expect to make quantitative inferences from models with complex interactions between market and non-market activities is a period of intuition-building. The CGE framework is made for this type of activity.

Beyond this, the marriage of CGE and techniques from non-market used in this experiment seem to confer ancillary benefits in a number of areas. The calibration procedure used in the CGE exercise incorporates information on unobserved heterogeneity in consumer tastes from the empirical model in a way that is not standard practice in the CGE literature, where individual-level variation is typically subsumed by the preference specification for a representative agent. Furthermore the solution techniques employed in contemporary CGE modeling appear to have promise as a strategy for addressing the challenges of welfare analysis in corner solution models as described above.

The evidence from the empirical exercise also supports the notion that general equilibrium effects may be important in our application. Similar to [30] we find negative effects of site congestion that appear to be substantial. While it is difficult to consider the congestion effect independently from other attributes it seems fair to extrapolate from our findings that congestion plays a role much larger than cross price effects, somewhat larger than direct water quality effects, and perhaps as great as own price effects in our model and application. These statements need to be conditioned, however, on the form of the congestion effect included. In using the proportion of people who visit a site as our measure of congestion we have ignored two potentially important determinants of that attribute: the intensity of use as well as the timing of use. Including these involves a re-parameterization of the model and transmission function, which are obvious directions for further research in both CGE and econometric settings.

The welfare calculations from the empirical model are sensible and intuitive and suggest that general equilibrium welfare measures can differ from their partial equilibrium counterparts in ways that have policy relevance. We again, however, must add caveats to this statement. The computational steps needed to compute both partial and general equilibrium welfare measures in this model are not yet fully understood and are the subject of ongoing research. Still, we have enough confidence in the results to conclude that feedback effects exist and for some counterfactual scenarios will cause divergence between the two welfare measures. The size and direction of the divergence will depend on the specifics of the scenario. This finding is also supported by the CGE modeling. Thus the evidence directly supported or implied by our modeling and results is that congestion matters, its effect is large enough to cause divergence between general and partial equilibrium welfare measures, and frameworks exist or are in development that can empirically measure the size and policy significance of these divergences.

In addition to this direct evidence several lessons and observations have emerged from the experience of carrying out this project. Perhaps most valuable is the appreciation gained of the combined technical and conceptual challenges that are inherent in addressing this type of problem. Our optimistic sense is that the computational challenges will not be limiting. The algorithms undeveloped at this stage are solvable either with brute computer force or clever programming and numerical analysis (probably in the end both). Perhaps more interesting will be the conceptual issues. We have found it challenging to rely on intuition to gauge the 'reasonableness' of results when feedback effects are present in the models. More to the point, it is not clear how to best carry out robustness and reality checks. For example, we discussed above that the transmission function used in our empirical analysis is restrictive in that it ignores the timing and intensity of site usage. How might we examine how robust our results are to this obvious simplification? More generally, how can researchers do sensitivity analysis on the numerous parametric stands that will need to be taken in order to specify a transmission function for the general equilibrium attribute? More subtly, how can we gauge the impact of 'limit effects' or the potential need for non-linear/discontinuous transmission functions? In our model, for example, closing a recreation site drives congestion for that site to zero. The zero congestion is mechanically transmitted to the rest of the model as an improvement in a site attribute – something that does not make intuitive sense as the result of a site closure. A less extreme examination of limit effects might be concerned with the range of exogenous data values that provide good endogenous attribute predictions via the transmission function versus those that extrapolate beyond the range of the function's ability to accurately predict. While this may seem a minor point it will be critical for gauging the accuracy of large-scale policy counterfactuals – exactly the type of policy scenarios for which we might expect re-equilibration to matter.

To these challenges we can add the fact that we have thus far only considered a simple sorting equilibrium application. Complex sorting equilibria add an additional layer of modeling to the mix, requiring that the analyst also take a stand on the form of the natural or other process that simultaneously (with behavior) determines the level of the endogenous attribute. Finally, we might wonder if multiple and interacting endogenous attributes are at play: for example, it may be that both congestion *and* fishing catch rates are endogenously determined and that these factors might interact in distributing peoples' choices. This notion presents conceptual, computational, and econometric challenges that could fill a career.

These challenges notwithstanding, our experience in this project has caused us to be optimistic about our ability to ultimately provide policy relevant inference on general equilibrium welfare measures arising from non-price equilibria. The work reported on represents progress in several technical areas and suggests avenues for further research in these areas. For example, the Lee and Pitt model as applied here is at the forefront of modeling the demand for quality differentiated goods. Its framework provides the potential for capturing rich parametric and stochastic substitution between commodities while allowing for non-additively separable preferences. Further research in this area could examine specifications that relax the independent errors assumption as well as increase the flexibility with which we characterize cross-price and income effects. The latter would involve a specification that moves away from the incomplete demand system approach in favor of modeling expenditures on outside goods as a substitute commodity – an approach that might also allow the inclusion of non-users in the model. Econometrically, the approach we have presented draws on three strands of literature: the classical use of Bayesian simulation methods, the true Bayesian econometric paradigm, and the tradition in empirical industrial organization of using two stages to deal with unobserved product attributes and endogenous attributes. Our exploitation of these ideas to date has been mainly informal and intuitive. There are obvious gains to a more careful (and formal) integration of these three econometric approaches in this class of problem.

X. References

1. Albert, J., and S. Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, **88**: 669-679.
2. Azevedo, C., K. J. Egan, J. A. Herriges, and C. L. Kling (2003) "The Iowa Lakes Valuation Project: Summary and Findings from Year 1," CARD Report, August.
3. Bayer, P., and R. McMillian, "Racial sorting and neighborhood quality," NBER working paper #11813, October 2006.
4. Bayer, P., R. McMillian, and K. Rueben (2005), "An equilibrium model of sorting in an urban housing market," NBER working paper #10865.
5. Bayer, P., N. Keohane, and C. Timmins (2006). "Migration and Hedonic Valuation: The Case of Air Quality." NBER Working Paper #12106.
6. Bayer, P. and C. Timmins (2005). "On the equilibrium properties of locational sorting models." *Journal of Urban Economics* 57: 462-477.
7. Berry, S.T. (1994). "Estimating discrete choice models of product differentiation," *Rand Journal of Economics* 25: 242-262.
8. Berry, S., J. Levinsohn, and A. Pakes (1995). "Automobile prices in equilibrium," *Econometrica* 63: 841-890.
9. Brooke A., D. Kendrick, and A. Meeraus (1992), *GAMS: A User's Guide*, Release 2.25, Scientific Press.
10. Carbone, J. C. and V. K. Smith (2007), "Evaluating Policy Interventions with General Equilibrium Externalities," *Journal of Public Economics*, forthcoming.
11. Epple, D., and H. Sieg (1999), "Estimating Equilibrium Models of Location Jurisdiction,"

Journal of Political Economy **107**(4): 645-681.

12. Ferreyra, Marie Marta (2006), "Estimating the Effects of Private School Vouchers in Multi-District Economies," *American Economic Review*, forthcoming.
13. Ferris, M. C. and T. S. Munson, (2007) PATH 4.6, http://www.gams.com/solvers/path.pdf.
14. Harrison, G. W. and H. D. Vinod, (1992) "The Sensitivity Analysis of Applied General Equilibrium Models: Completely Randomized Factorial Sampling Designs," *The Review of Economics and Statistics*, pp. 357-362.
15. Harrison, G. W., R. Jones, L. J. Kimbell and R. Wigle (1993), "How Robust is Applied General Equilibrium Analysis?" *Journal of Policy Modeling*, pp. 99-115.
16. LaFrance, J. T., (1985), "Linear Demand Functions in Theory and Practice," *Journal of Economic Theory*, **37**(1): 147-166.
17. Lee, L. F., and M. M. Pitt (1986). "Microeconomic Demand Systems with Binding Non-Negativity Constraint: the Dual Approach," *Econometrica* 54: 1237-1242.
18. Manski, C. F. (2000). "Economic analysis of social interactions," *Journal of Economic Perspectives* 14: 115-136.
19. Massey, D. M., S. C. Newbold, and B. Gentner (2006), "Valuing Water Quality Changes Using a Bioeconomic Model of a Coastal Recreational Fishery," *Journal of Environmental Economics and Management* **52**(1): 482-500.
20. Murdock, J., (2006), "Handling unobserved site characteristics in random utility models of recreation demand," *Journal of Environmental Economics and Management* 51: 1-25.
21. Orhun, A.Y. (2005). "Spatial differentiation in the supermarket industry," working paper, Haas School of Business, University of California Berkeley.
22. Phaneuf, D.J. (1999). "A Dual Approach to Modeling Comer Solutions in Recreation

Demand." *Journal of Environmental Economics and Management* 37 (1), 85-105.

23. Phaneuf, D.J., Kling, C.L., Herriges, J.A. (2000). "Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand." *Review of Economics and Statistics* 82 (2), 83-92.
24. Pitt, M., and D. L. Millimet (2003), "Estimation of Coherent Demand Systems with Many Non-Negativity Constraints," Working Paper, October.
25. Seim, K. (2005). "An empirical model of firm entry with endogenous product type choices," working paper, Stanford Graduate School of Business.
26. Schelling, T.C., (1978), *Micromotives and Macrobehavior,* New York: Norton.
27. Smith, M.D., J. N. Sanchirico, and J.E. Wilen (2007). "The economics of spatial-dynamic process: application to renewable resources," RFF discussion paper.
28. Smith, M.D., J. Zhang, and F.C. Coleman (2007). "Economic modeling of fisheries with complex life histories: Avoiding biological management failures," working paper, Duke University.
29. Smith, V.K, Sieg, H., Banzhaf, H.S., Walsh, R. (2004). "General Equilibrium Benefits for Environmental Improvements: Projected Ozone Reduction Under EPA's Prospective Analysis For The Los Angeles Air Basin". *Journal of Environmental Economics and Management* 47, 559-584.
30. Timmins, C., and J. Murdock (2007) "A Revealed Preference Approach to the Measurement of Congestion in Travel Cost Models." *Journal of Environmental Economics and Management* 53: 230-249.
31. von Haefen, R. and D.J. Phaneuf (2005). "Kuhn-Tucker Demand System Approaches to Non-market Valuation", in *Applications of Simulation Methods in Environmental and*

Resource Economics, A. Alberini and R. Scarpa, eds., Springer.

32. von Haefen, R., D. Phaneuf, and G. Parsons (2004) "Estimation and Welfare Analysis with Large Demand Systems," *Journal of Business and Economic Statistics* 22(2): 192-205.
33. Walsh, R., "Endogenous open space amenities in a locational equilibrium," *Journal of Urban Economics* 61 (2007): 319-344.
34. Wang, Y. (2003), "Estimation of Recreation Demand Models Using Dual Approaches," Ph.D. Dissertation, North Carolina State University.
35. Yang, S., Y. Chen, and G. Allenby (2003). "Bayesian analysis of simultaneous demand and supply," *Quantitative Marketing and Economics* 1: 251-275.

Table 1: Summary of Selected Parameters from Posterior Distributions: 1st Stage[a]

Parameter	Posterior Mean	Posterior Std. Deviation	Mean/Std. Deviation	Posterior Median
β_{own}	-0.046	0.0015	-30.596	-0.0459
β_{cross}	0.0002	0.000	20.10	0.0002

[a]Calculated using 800 simulated draws from the posterior distribution. Posterior summaries for intercepts and variances shown in a subsequent table.

Table 2: Summary of Selected Parameters from Posterior Distributions: 2nd Stage OLS Decomposition[a]

Parameter	Posterior Mean	Posterior Std. Deviation	Mean/Std. Deviation	Posterior Median
γ_0	-5.435	0.2015	-26.96	-5.429
γ_{secchi}	0.0391	0.0862	0.453	0.0411
$\gamma_{chlorophyll}$	0.0013	0.0017	0.744	0.0014
$\gamma_{congestion}$	-13.34	2.14	-6.22	-13.32

[a]Calculated via OLS regressions of each first stage draw of the intercepts on the site characteristics. Summaries are calculated from the resulting 800 sets of 2nd stage estimates.

Table 3: Summary of Selected Parameters from Posterior Distributions: 2nd Stage IV Decomposition[a]

Parameter	Posterior Mean	Posterior Std. Deviation	Mean/Std. Deviation	Posterior Median
γ_0	-4.081	0.2225	-18.34	-4.072
γ_{secchi}	0.1122	0.0867	1.294	0.1128
$\gamma_{chlorophyll}$	-0.0064	0.0019	-3.36	-0.0064
$\gamma_{congestion}$	-55.91	4.29	-13.00	-55.78

[a]Calculated via IV regressions of each first stage draw of the intercepts on the site characteristics. Summaries are calculated from the resulting 800 sets of 2nd stage estimates.

Table 4: Point Estimates for Welfare Effects[a]

Counterfactual Scenario	Partial Equilibrium Estimate		General Equilibrium Estimate	
	Sample mean	Sample Median	Sample Mean	Sample Median
Scenario 1: Loss of nine highly popular sites.	-$630.40	-$67.64	-$726.11	-$142.84
Scenario 2: Loss of nine moderately popular sites	-$51.86	$0.00	-$58.21	-$5.08
Scenario 3: Widespread small quality improvements	$207.92	$146.01	$87.28	$61.16
Scenario 4: Localized large quality improvements	$50.23	$17.14	$21.09	$7.57

[a]Measured in dollars per active lake visitor person per year

Table 5: CGE Welfare Estimates – Totals Specification

	Half Congestion Effect			Full Congestion Effect			Double Congestion Effect		
Scenario	PE	GE	%Diff	PE	GE	%Diff	PE	GE	%Diff
1	-552.3	-549.1	-0.6	-524.5	-527.2	0.5	-469.0	-506.7	7.4
2	-195.4	-193.3	-1.1	-193.3	-189.9	-1.8	-189.0	-184.2	-2.6
3	157.7	126.2	-25.0	153.6	106.7	-44.0	145.3	88.1	-64.9
4	59.7	49.4	-20.9	58.2	41.8	-39.2	55.2	32.8	-68.1

Table 6: CGE Welfare Estimates – Totals vs. Shares Specifications

	Totals Specification			Shares Specification		
Scenario	PE	GE	%Diff	PE	GE	%Diff
1	-524.5	-527.2	0.5	-580.0	-618.3	6.2
2	-193.3	-189.9	-1.8	-197.6	-212.1	6.9
3	153.6	106.7	-44.0	161.9	172.2	6.0
4	58.2	41.8	-39.2	61.2	56.0	-9.2

[1] Recent notable exceptions are the equilibrium sorting models of [29], [30], and [5].

[2] A similar approach was employed by Wang [34] and Pitt and Millimet [24].

[3] In practice it is also necessary to check that the regimes implied by $\tilde{z}_{ij}^{*t}$ are in the feasible set and, if not, adjust using the strategy discussed above. This emphasizes the point that using predicted behavior to simulate new endogenous attribute outcomes also depends critically on the ability to solve consumers' outcomes accurately and quickly.

[4] Appendix Figure B.1 provides histograms of the full marginal posterior distributions for the price parameters in Table 1 and on the attributes for the specification in Table 3.

[5] Appendix Table B.1 provides the posterior means and standard deviations for the sites-specific intercepts and variance terms.

[6] It is worth noting that, for this logic to be a complete description of the equilibrating process, it must not be the case that this adjustment induces changes in the virtual prices at zero-visit sites. If there were systematic changes in these prices across individuals in response to the policy, this could complicate the equilibrium adjustment story described here.

Appendix A. Estimation Algorithm

Our estimation algorithm proceeds in two steps. The first stage uses a Gibbs sampler to accumulate draws from the posterior distribution for $\theta=(\alpha_1,...,\alpha_J, \beta_1, \beta_2, \sigma_1,...,\sigma_J)$. The second stage uses linear regression to decompose the intercepts $\alpha_1,...,\alpha_J$ into observed and unobserved components. In this appendix we provide additional technical details on the two steps used in estimation.

First Stage Estimation

The following steps for the first stage are involved in a simulation process drawing M values from the posterior distribution:

Step 1: Set starting values

Initial values $\psi(0)$ and $\sigma_j^2(0)$ for the simulation are established. For example, one might obtain starting values in the current context by running a simple tobit model for each site using observed trips data, yielding site specific intercepts and variance terms (i.e., α_j and σ_j^2) and, by averaging across sites, starting values for the own- and cross-price parameters.

Step 2: Drawing posterior notional demands from $p(z^*|z,\psi,\Sigma)$

The notional demands, conditional on observed demands and the parameters of the model, are drawn from a truncated normal distribution for non-consumed goods and from a normal distribution for the consumed goods. Specifically, from equation (16) in the paper, we have that

$$\tilde{\varepsilon}_{ij}(m) = z_{ij} - \tilde{\alpha}_j(m-1) - \tilde{\beta}_{jC}(m-1)\, p_{iC}, \quad j = r+1,\dots,J, \tag{A.1}$$

where the parenthetical arguments denote the step in the iteration process (m=1,…,M). Since $\tilde{\varepsilon}_{iC} = \varepsilon_{iC} - \beta_{CN}\beta_{NN}^{-1}\varepsilon_{iN}$, then $\left(\varepsilon'_{iN}, \tilde{\varepsilon}'_{iC}\right)' \sim N(0,\Omega)$ where

$$\Omega = Cov\begin{bmatrix} \varepsilon_{i1}\varepsilon_{i1} & \cdots & \varepsilon_{i1}\varepsilon_{ir} & \varepsilon_{i1}\tilde{\varepsilon}_{i,r+1} & \cdots & \varepsilon_{i1}\tilde{\varepsilon}_{iJ} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{i1}\varepsilon_{ir} & \cdots & \varepsilon_{ir}\varepsilon_{ir} & \varepsilon_{ir}\tilde{\varepsilon}_{i,r+1} & \cdots & \varepsilon_{ir}\tilde{\varepsilon}_{iJ} \\ \varepsilon_{i1}\tilde{\varepsilon}_{i,r+1} & \cdots & \varepsilon_{ir}\tilde{\varepsilon}_{i,r+1} & \tilde{\varepsilon}_{i,r+1}\tilde{\varepsilon}_{i,r+1} & \cdots & \tilde{\varepsilon}_{iJ}\tilde{\varepsilon}_{i,r+1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{i1}\tilde{\varepsilon}_{iJ} & \cdots & \varepsilon_{ir}\tilde{\varepsilon}_{iJ} & \tilde{\varepsilon}_{i,r+1}\tilde{\varepsilon}_{iJ} & \cdots & \tilde{\varepsilon}_{iJ}\tilde{\varepsilon}_{iJ} \end{bmatrix} \tag{A.2}$$

$$= \begin{bmatrix} \Sigma_{NN} & \Sigma_{NC} - \Sigma_{NN}\beta_{NN}^{-1}\beta_{NC} \\ \Sigma_{CN} - \beta_{CN}\beta_{NN}^{-1}\Sigma_{NN} & \Sigma_{CC} + \beta_{CN}\beta_{NN}^{-1}\Sigma_{NN}\beta_{NN}^{-1}\beta_{NC} - \beta_{CN}\beta_{NN}^{-1}\Sigma_{NC} - \Sigma_{CN}\beta_{NN}^{-1}\beta_{NC} \end{bmatrix}.$$

In the current application, this expression reduces considerably, since $\Sigma_{NC}=0_{N\times C}$, so that:

$$\Omega = \begin{bmatrix} \Sigma_{NN} & -\Sigma_{NN}\beta_{NN}^{-1}\beta_{NC} \\ -\beta_{CN}\beta_{NN}^{-1}\Sigma_{NN} & \Sigma_{CC} + \beta_{CN}\beta_{NN}^{-1}\Sigma_{NN}\beta_{NN}^{-1}\beta_{NC} \end{bmatrix}. \tag{A.3}$$

The notional demands for the non-consumed goods correspond to $z_{ij}^* < 0$, or equivalently:

$$\varepsilon_{ij} < -\left[\alpha_j + \sum_{j=1}^{J}\beta_{jk}p_{ik}\right] \equiv \delta_{ij}, \quad j = 1,\ldots,r. \tag{A.4}$$

Using the expression in (A.4), we sequentially obtain $\varepsilon_{ij}(m) \mid \varepsilon_{i,-j}(m)$, where

$\varepsilon_{i,-j}(m) = \left(\varepsilon_{i,1}(m),\ldots,\varepsilon_{i,j-1}(m),\varepsilon_{i,j+1}(m-1),\ldots,\varepsilon_{i,r}(m),\tilde{\varepsilon}_{iC}(m)\right)$ denotes all of the errors for individual *i* except $\varepsilon_{ij}(m)$, noting that

$$\varepsilon_{ij} \mid \varepsilon_{i,-j} \sim TN_{(-\infty,\delta_{ij})}\left(\mu_{ij|-j},\omega_{j|-j}\right), \quad j = 1,\ldots,r, \tag{A.5}$$

where

$$\mu_{j|-j} = \Omega_{j,-j}\Omega_{-j,-j}^{-1}\varepsilon_{i,-j}, \quad j = 1,\ldots,r, \tag{A.6}$$

and

$$\omega_{j|-j} = \omega_{jj} - \Omega_{j,-j}\Omega_{-j,-j}^{-1}\Omega_{-j,j}, \quad j = 1,\ldots,r. \tag{A.7}$$

Draws of $\varepsilon_{ij}(m)$ for j=1,…,r are obtained using (A.5) and the inversion method. Values of $\varepsilon_{ij}(m)$ for j=r+1,…,J are then obtained using

$$\varepsilon_{ij}(m) = \tilde{\varepsilon}_{ij}(m) + \beta_{CN}\beta_{NN}^{-1}\varepsilon_{iN}(m), \quad j = r+1,\ldots,J. \tag{A.8}$$

Finally, the notional demands are formed using

$$z_{ij}^{*}(m) = \alpha_j(m-1) + \sum_{j=1}^{J}\beta_{jk}(m-1)\,p_{ik} + \varepsilon_{ij}(m), \quad j = 1,\ldots,J. \tag{A.9}$$

Step 3: Drawing posterior covariance terms from $p(\Sigma|z^{*},z,\psi)$

The variance terms σ_j^2 's are updated using the conditional posterior distributions

$\sigma_j^2 \mid z_j^{*},\psi \sim IG\left(1+N,\bar{s}_j\right)$, where $\bar{s}_j = \left(\underline{s}_j + Ns_j^2\right)\big/\left(1+N\right)$ and s_j^2 denotes the sampling variance

from the ordinary least squares estimation of

$$z_{ij}^{*}(m) = \alpha_j(m) + \sum_{j=1}^{J}\beta_{jk}(m)\,p_{ik}, \quad j = 1,\ldots,J. \tag{A.10}$$

The updated value of $\sigma_i^2(m)$ is drawn from $IG\left(1+N,\bar{s}_j(m)\right)$.

Step 4: Drawing posterior parameters from $p(\psi|z^{*},z,\Sigma)$

Conditional on the notional demands and the variance-covariance structure, the posterior distribution of the parameters $\psi=(\alpha',\beta_1,\beta_2)$ is normally distributed. Specifically,

$$\psi \sim N\left(\hat{\psi},\hat{\Sigma}_{\psi}\right), \tag{A.11}$$

where $\hat{\psi}$ denotes the SUR estimates of ψ and $\hat{\Sigma}_{\psi}$ denotes the corresponding covariance matrix of the estimated parameters. The parameter vector $\psi(m)$ is then obtained by drawing from $N\left(\hat{\psi}(m),\hat{\Sigma}_{\psi}(m)\right)$.

Steps 2 through 4 are repeated for m=1,…,M. The first B draws are used for burn-in and discarded. From the remaining M-B draws, every fourth draw is retained, yielding a total of (M-B)/4 draws from the posterior distribution of interest. In our application, we use M=7200 and B=*4000*, yielding a total of 800 draws for the posterior distribution.

Second Stage Estimation

Identification of the effects of observable site attributes on demand is feasible if two conditions are met. First, J must be large enough and there must be enough variation over the q_j's to provide sufficiently precise estimates from the linear model employed. In our application J=128, which we find to be large enough to estimate a small number of site attribute effects. Second, the observable attributes q_j must be uncorrelated with the unobserved attributes ξ_j or instruments must be available for the (econometrically) endogenous explanatory variables. In our application this includes congestion defined as the proportion of people in the user population who visit a particular site, and an instrument is needed for this variable.

We construct an instrument by estimating a binary logit model for each site, where the dependent variable is equal to one if the person visits the site and zero otherwise. The logit probability is parameterized to include all the attributes of the site thought to be exogenous, such as travel costs and ambient water quality, and the parameter estimates used to construct predictions for each person's probability of visiting each site. In practice we estimate the binary outcomes for all sites jointly and constrain the parameters on the explanatory variables to be the same across all sites. We define our instrument for s_j^0 to be the average of the predictions for site j for all people in the sample. We find this measure to be reasonably well correlated with s_j^0 and plausibly exogenous since it is functionally related only to exogenous variables. As noted in [30], the power of this instrument strategy depends on the degree to which the exogenous site attributes are good predictors of visitation shares.

Appendix B: Additional Estimation Results

Table B.1: Posterior Summary for Intercepts and Variance Parameters

Site	Posterior Mean α_j	Posterior std. Deviation α_j	Posterior Mean σ_j^2	Posterior Std. Deviation σ_j^2
site1	-6.853	0.5698	21.6205	2.8972
site2	-3.9804	0.5177	16.5106	2.4113
site3	-3.5416	0.472	6.5615	1.2099
site4	-4.5809	0.5179	7.9619	1.3855
site5	-7.1232	0.6516	25.9651	3.5893
site6	-7.4964	0.7112	49.7431	5.6995
site7	-4.6257	0.4944	8.6021	1.4533
site8	-7.5566	0.7139	43.9953	5.5753
site9	-7.0336	0.721	98.8242	10.8453
site10	-4.0317	0.7095	64.0104	7.6369
site11	-7.2933	0.7324	49.9359	6.0285
site12	-3.9824	0.5992	19.1432	2.7964
site13	-3.9182	0.5773	11.292	1.9846
site14	-6.643	0.6569	26.8866	3.6205
site15	-3.5827	0.5861	21.062	3.005
site16	-7.0708	0.6815	56.1644	6.6652
site17	-4.594	0.6589	26.8254	3.8257
site18	-7.2464	0.6727	26.7819	3.6761
site19	-3.5336	0.5347	16.9989	2.4502
site20	-5.8983	0.5825	17.0884	2.6514
site21	-6.4162	0.7842	124.3539	13.5503
site22	-4.1526	0.6058	18.1337	2.8325
site23	-5.2395	0.7264	116.3501	12.4506
site24	-3.4916	0.4825	10.2819	1.6842
site25	-5.0832	0.6614	22.852	3.4256
site26	-6.3228	0.5336	19.2275	2.6721
site27	-4.5309	0.6175	31.509	4.0606
site28	-7.1968	0.6494	27.5623	3.7272
site29	-3.6768	0.5184	9.8924	1.8605
site30	-6.8934	0.6903	32.5796	4.5013
site31	-6.1151	0.6361	21.9603	3.3361
site32	-3.5645	0.6951	74.1869	8.4062
site33	-7.0511	0.7023	37.253	5.6784
site34	-3.143	0.3565	3.0317	0.6714
site35	-5.7108	0.667	33.1604	4.24
site36	-2.5374	0.369	4.0039	0.7492
site37	-7.7321	0.6936	52.9631	6.14
site38	-4.6876	0.4236	5.2954	0.9961
site39	-6.2413	0.6191	15.0001	2.5055
site40	-5.7511	0.6308	27.362	3.7193
site41	-5.5443	0.5887	16.054	2.2652
site42	-6.9314	0.6694	21.339	3.3877
site43	-7.01	0.6769	26.6037	3.9212
site44	-8.4337	0.7326	41.1814	5.4627
site45	-3.8146	0.3784	3.303	0.7044
site46	-5.4035	0.6383	23.6429	3.6075
site47	-3.9198	0.5969	18.939	2.8882
site48	-7.5834	0.6577	37.7102	4.9194
site49	-7.3969	0.7128	39.0452	5.3381
site50	-6.0868	0.6992	28.6422	4.2218
site51	-6.0373	0.6134	25.8125	3.5502
site52	-6.5316	0.6692	35.1292	4.7759
site53	-5.8525	0.71	40.6978	5.4499
site54	-4.6094	0.6228	15.7547	2.4939
site55	-5.317	0.6624	28.9138	4.0891

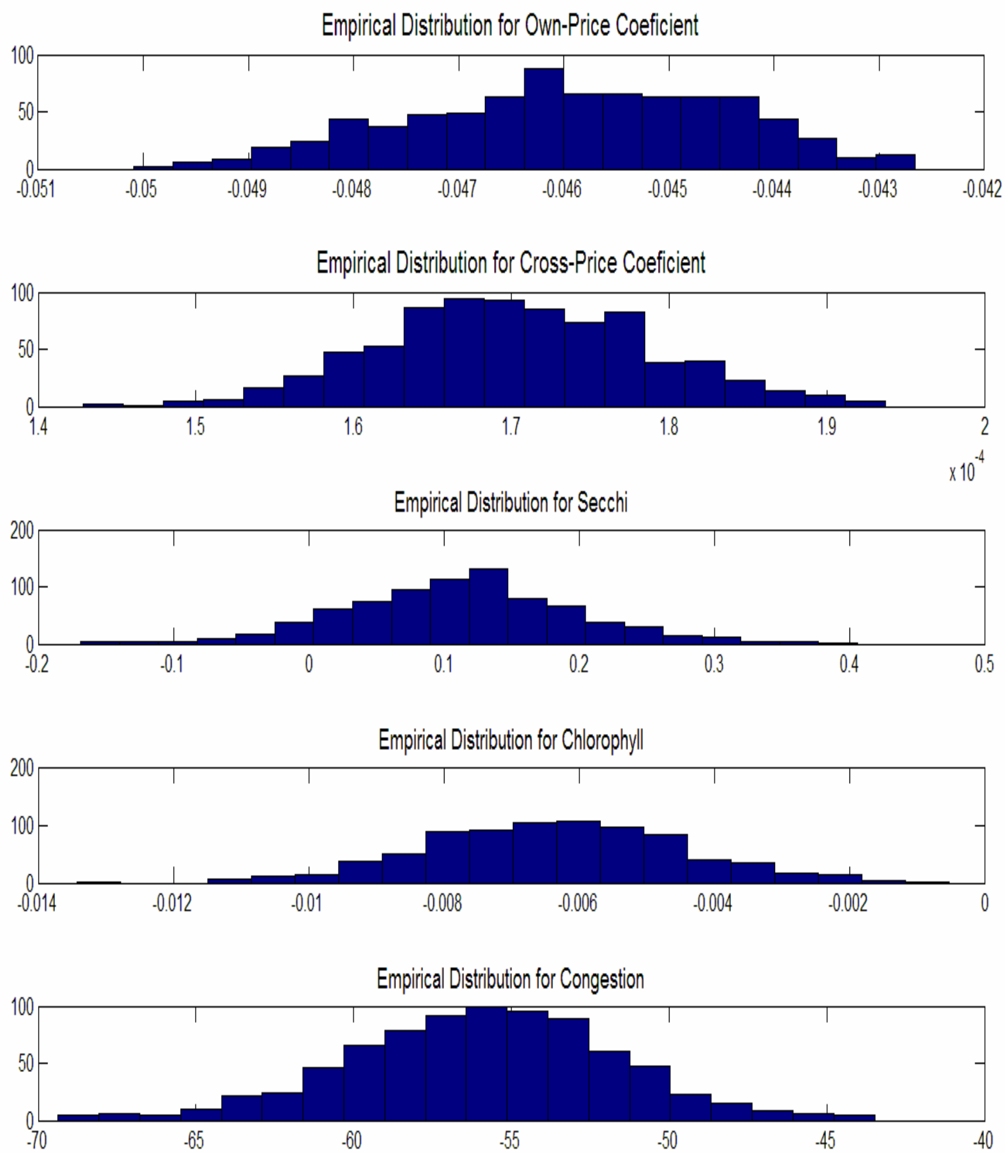
Table B.1: Posterior Summary for Intercepts and Variance Parameters

Site	Posterior Mean α_j	Posterior std. Deviation α_j	Posterior Mean σ_j^2	Posterior Std. Deviation σ_j^2
site56	-7.4607	0.6787	26.8651	3.8537
site57	-6.7462	0.6406	25.4104	3.5819
site58	-6.8167	0.8778	69.6435	8.7107
site59	-6.6254	0.7072	92.0742	10.3736
site60	-7.13	0.6678	31.5164	4.32
site61	-4.0556	0.6159	18.9596	2.9636
site62	-7.1759	0.7303	38.1587	5.3751
site63	-5.6861	0.8047	37.9029	5.6098
site64	-3.1248	0.476	4.2605	0.981
site65	-2.4601	0.5749	14.8261	2.3857
site66	-4.2242	0.5623	11.2038	2.0147
site67	-6.3264	0.7774	41.5089	5.939
site68	-6.4949	0.8175	38.582	5.652
site69	-5.6505	0.7156	21.5728	3.3974
site70	-3.728	0.5263	11.1625	1.9554
site71	-4.483	0.7614	44.4682	6.0189
site72	-7.1417	0.687	29.4436	4.2063
site73	-4.5442	0.551	10.333	1.7792
site74	-5.2364	0.678	32.8502	4.4143
site75	-5.0529	0.7638	33.7347	5.1076
site76	-8.4861	0.8132	44.8233	6.2416
site77	-3.2839	0.4585	4.7608	1.0382
site78	-7.1689	0.6732	20.3329	3.2558
site79	-2.5307	0.4015	1.7805	0.4951
site80	-6.0551	0.553	12.0678	1.9475
site81	-3.7554	0.5977	11.5057	2.2217
site82	-5.5906	0.5478	9.5444	1.7068
site83	-4.1143	0.5724	9.1745	1.7158
site84	-5.1554	0.6426	14.0223	2.3759
site85	-3.8837	0.5858	10.5073	2.009
site86	-2.7552	0.535	4.9138	1.1047
site87	-6.5936	0.7421	35.0382	4.8049
site88	-3.7051	0.5349	8.819	1.6471
site89	-7.3492	0.7675	25.5976	4.1353
site90	-7.8308	0.7873	44.3108	6.075
site91	-3.7418	0.6025	12.2716	2.0722
site92	-7.8576	0.8142	63.3195	8.7261
site93	-6.728	0.8084	30.2937	4.7713
site94	-6.9139	0.7567	36.3674	5.1407
site95	-8.0552	0.8184	110.3577	11.8634
site96	-7.7719	0.7616	33.6603	4.7278
site97	-7.8101	0.8437	115.6849	13.7802
site98	-7.7368	0.7068	27.2369	4.0248
site99	-8.8789	0.8622	50.4859	7.1706
site100	-6.029	0.5733	8.7578	1.6529
site101	-5.7964	0.7613	135.7294	15.2796
site102	-5.7679	0.664	13.4248	2.5549
site103	-3.9927	0.7187	23.3549	3.8011
site104	-4.9423	0.6581	18.629	3.0432
site105	-5.0798	0.7101	17.3794	3.1948
site106	-6.2556	0.6396	13.7435	2.4165
site107	-6.3561	0.6834	17.4409	2.8836
site108	-6.0069	0.8007	18.8009	3.4047
site109	-7.1128	0.8412	65.8239	8.657
site110	-7.3213	0.7809	35.4324	5.141
site111	-4.6368	0.5793	8.9574	1.75
site112	-7.9943	0.7794	51.582	6.5441
site113	-4.1814	0.6463	10.4926	2.0675

Table B.1: Posterior Summary for Intercepts and Variance Parameters

Site	Posterior Mean α_j	Posterior std. Deviation α_j	Posterior Mean σ_j^2	Posterior Std. Deviation σ_j^2
site114	-2.5684	0.467	5.0929	1.0555
site115	-7.9598	0.8528	51.4775	6.9238
site116	-8.6359	0.9268	37.511	6.1373
site117	-5.28	0.7707	29.4527	4.494
site118	-9.9038	0.9543	64.0853	8.7494
site119	-5.6115	0.7516	30.5258	4.6671
site120	-7.2441	0.8385	47.4747	6.4673
site121	-4.7558	0.9157	98.7864	12.4806
site122	-7.4532	0.7654	25.5415	3.7944
site123	-5.3224	0.5757	9.008	1.5731
site124	-3.5836	0.541	3.6977	0.9531
site125	-4.1893	0.6111	11.6769	2.0662
site126	-3.9118	0.6452	8.4505	1.7459
site127	-4.6318	0.7037	13.6563	2.4983
site128	-5.0304	0.6462	18.0508	2.8207

Figure B.1: Histograms Showing Empirical Distributions



Appendix C: Calibration of a Random Parameters CGE Model

A basic premise of the estimation strategy is that there are unobserved components of individual tastes due to the fact that non-negativity constraints on demands are binding at many of the sites in the benchmark dataset. Because they typically use aggregate datasets, CGE applications rarely attempt to take this type of heterogeneity into account.[1] What follows is the description of a method for incorporating the information in the posterior distributions from the econometric model into account in the calibration of the Iowa Lakes CGE application.

The specific form that the demand functions take in the econometric model is:

$$z_{ij} = \gamma_0 + \xi_j + \sum_m \gamma_m q_{mj} + \sum_k \beta_{jk} \pi_{ik} + \varepsilon_{ij} \geq 0 \;\perp\; \pi_{ij} \leq p_{ij} \qquad \text{(C.1)}$$

where the difference between (C.1) and (27) is the explicit addition of an error term ε_{ij} and site-specific unobserved attribute ξ_j in (C.1). In order to calibrate the simulation model in a manner consistent with this logic, we must allow for the fact that $\gamma_{ij}^0 = \gamma_0 + \xi_j + \varepsilon_{ij}$ may take on a range of values that are consistent with replicating the benchmark dataset. To see this, suppose we observe a corner solution at site *j* for individual *i* in the benchmark data. It may be the case that the non-negativity constraint is weakly binding so that a marginal reduction in the travel cost to that site would induce visits. The alternative is that it is strongly binding in which case even a substantial increase in the attractiveness of the site might yield no increase in visitation. Simulating policy experiments and calculating the welfare consequences under these two different scenarios would obviously lead to different results at the individual level and, if these differences are systematic across individuals, at the aggregate level as well.

[1] Notable exceptions include [14, 15].

Our approach to calibrating these γ_{ij}^{0} parameters is to take draws from the posterior distribution of these parameters described by the estimation results conditional on matching the benchmark demands in the data. By repeating our counterfactual experiments for different sets of draws from these distributions, we can describe the distribution of the outcome measures that we are interested in such as equilibrium visitation and congestion levels at different sites and welfare changes for different types of individuals. This parallels the logic used in the counterfactual analysis based on the empirical model.

While there is not a unique mapping of the γ_{ij}^{0} terms to observed demands, the conceptual logic of the virtual prices described above does impose some structure on the calibration problem. Specifically, it pins down the value of the virtual price so that $\pi_{ij}=p_{ij}$ if the benchmark number of visits that an individual takes to site j is non-zero. Thus

$$\gamma_{ij}^{0} = \tilde{z}_{ij} - \left[\sum_{m} \gamma_m \tilde{q}_{mj} + \beta_{jj} p_{ij} + \sum_{k \neq j} \beta_{jk} \pi_{ik}\right] \geq 0 \quad \forall j \;\; s.t. \;\; \tilde{z}_{ij} > 0, \tag{C.2}$$

where $\tilde{z}_{ij}$ and $\tilde{q}_{mj}$ are the benchmark levels of site visits and attribute levels observed in the data for each site.

For site demands that are zero in the benchmark data, we draw the γ_{ij}^{0} randomly according to the posterior distribution for these terms. These draws imply specific realizations of the virtual prices associated with each demand because they must replicate the observed demand of zero given γ_{ij}^{0} and the other virtual prices in the demand system. Thus,

$$\pi_{ij} = -\left[\gamma_{ij}^{0} + \sum_{m} \gamma_m \tilde{q}_{mj} + \sum_{k \neq j} \beta_{jk} \pi_{ik}\right] / \beta_{jj} \geq 0 \quad \forall j \;\; s.t. \;\; \tilde{z}_{ij} = 0. \tag{C.3}$$

The calibration of the free γ_{ij}^{0} parameters (those for the positive demand sites) and the virtual prices (i.e., those for the zero-demand sites) are characterized by the solution to the system of equations defined by (C.2) and (C.3). For a given individual, if N is the number of non-zero sites demands and Z is the number of zero site demands, this strategy produces a system of $N+Z$ equations and $N+Z$ unknowns – N of the γ_{ij}^{0} terms and Z of the π_{ij} terms.

The solution to this system of equations is not, however, guaranteed to produce a set of virtual prices that do not violate the condition of the model that these prices be less than the observed travel cost prices to each site. Therefore, the full calibration routine involves sequentially solving this system of equations, checking to insure that the resulting virtual prices do not violate their upper bounds, taking new draws from the γ_{ij}^{0} distributions for those demands that do not violate this condition, and solving for the new virtual prices until no such violations occur.