

Combining Aggregate Demand and Discrete Choice Data with Application to Deer License Demand in Indiana

Carson Reeling *Associate Professor, Department of Agricultural Economics, Purdue University, West Lafayette, Indiana; creeling@purdue.edu*

Dane Erickson *Undergraduate Research Assistant, Department of Agricultural Economics, Purdue University, West Lafayette, Indiana; erick310@msu.edu*

Yusun Kim *Graduate Research Assistant, Department of Agricultural Economics, Purdue University, West Lafayette, Indiana; yousun0129@gmail.com*

John G. Lee *Professor, Department of Agricultural Economics, Purdue University, West Lafayette, Indiana; jlee1@purdue.edu*

Nicole J.O. Widmar *Professor, Department of Agricultural Economics, Purdue University, West Lafayette, Indiana; nwidmar@purdue.edu*

ABSTRACT Estimating demand for licenses for recreational activities is complicated due to a lack of meaningful variation across time, space, buyer types, and license attributes, including price. Prior work uses discrete choice experiments (DCEs) to overcome this challenge, but the resulting demand models are unlikely to replicate observed demands in the absence of ad hoc calibration procedures. We use a generalized method of moments-based approach that combines

DCE data with observed market share data to estimate a choice model that yields demand functions that much more closely replicate observed data. (JEL Q21, Q26)

1. Introduction

Resource management agencies in the United States are tasked with setting license prices and other attributes like seasons, equipment restrictions and, in the case of hunting and fishing licenses, daily and seasonal bag limits and quotas. Their objectives in making choices about licenses offerings may include maximizing agency revenues, conserving or managing scarce natural resources, and encouraging participation among constituents. Understanding how license prices and other attributes influence license demand is important for making efficient resource management decisions.

Estimating demand for recreational licenses can be challenging. Some prior work estimates demand for recreational licenses using aggregate sales data by exploiting variation in the nominal price of licenses over time due to inflation (Loomis et al. 2000; Erickson et al. 2019). However, these prior approaches are not useful for estimating demand for license attributes, which can be important when resource management agencies seek to change these attributes—say, by increasing license prices or changing bag limits or equipment restrictions—or change the structure of license offerings altogether. Revealed preference approaches that could elicit this information (e.g., choice modeling or hedonic analysis using micro-level license sales data) are typically inappropriate since there is often no useful variation in prices or other attributes with which to identify preference parameters. Indeed, license prices and attributes are typically fixed by statute, are uniform across users (although there may be some coarse price

discrimination by age, residency, or landownership status), and tend not to vary for years at a time.

Stated preference approaches—particularly discrete choice experiments (DCEs; Train 2009)—can overcome these challenges. DCEs work by presenting subjects with descriptions of different hypothetical products that vary in their attributes, including price. Subjects are asked to choose which if any of the hypothetical products they would purchase from a given set of alternatives. From this information, analysts can estimate the probability different subjects would choose a particular product, which can be extrapolated to infer market shares. Further, analysts can estimate preferences for product attributes by observing how subjects trade off these attributes for price. Because DCEs allow analysts to control variation in prices and other product attributes, they can be useful in estimating preference parameters for goods like recreation licenses for which attributes are effectively fixed over time and space. Several prior studies use DCEs to estimate hunter demand for hunting license attributes for big game species including white-tailed deer (Mackenzie 1990; Serenari et al. 2019) and turkey (Schroeder et al. 2018) as well as attributes of hunting clubs (Mingie et al. 2017).

Demand estimated from DCEs is unlikely to replicate observed demand in aggregate, and this is well-known (e.g., Vossler et al. 2012). Aside from hypothetical bias and other complications, this is because respondents to DCEs are typically asked to choose between just a few alternatives, and the attributes of the alternatives are selected for statistical efficiency rather than to perfectly match real-world offerings. This lack of replication can be problematic when performing counterfactual analysis using these estimated models—say, to simulate the effect of regulatory changes to license structure. Indeed, the validity of any simulated result is questionable when the simulated baseline does not match reality.

One solution to this replication problem is *ex post* calibration of the estimated preference parameters (Louviere et al. 2000). A common approach to calibration involves adding constant terms to the utility functions estimated from DCE data, then calculating the values of these terms that result in choice probabilities that exactly replicate observed market shares (e.g., Naidoo and Adamowicz 2005). While this calibration procedure is elegant, it seems a bit incongruous to simulate behavior using different utility functions than one estimated. The availability of information about aggregate observed behavior also raises the question of whether this information can be more profitably used to improve estimation of individual-level behavioral models in the first place.

To this end, Imbens and Lancaster (1994) develop an estimation approach that combines aggregate and individual-level data to more efficiently estimate microeconomic models. Briefly, their approach involves specifying a micro-level behavioral model, then deriving moment conditions equating predicted behavior from the model to known moments of the population distribution from aggregate data (e.g., from national statistics). Behavioral model parameters can then be estimated from these conditions using the generalized method of moments (GMM). Critically, and to the extent that these moment conditions hold, the estimated micro model should imply individual behaviors that aggregate to the observed, population-level behavior, obviating the need for calibration or other *ad hoc* procedures for ensuring replicability.

We follow Imbens and Lancaster's (1994) approach by combining aggregate license sales data with our DCE data to estimate demand for white-tailed deer hunting licenses and their attributes in Indiana. Compared to those from standard choice modeling approaches, our estimated demand functions much more closely replicate observed demand for existing deer licenses. Further, our parameters are estimated more efficiently such that predictions from

counterfactual analyses using our GMM-based model (e.g., changes to license demand or hunter surplus following changes to license structure) are more precise. We are not aware of prior work that applies this approach to estimating choice models from DCEs.

Our work is related to the extensive literature on combining stated and revealed preference data for estimating willingness to pay for nonmarket goods (see Whitehead et al. 2008 for a review). In particular, it is closely related to the study by Phaneuf et al. (2013), who estimate marginal willingness to pay for distance from hazardous waste sites in residential location choices. The authors use home sales data to estimate a hedonic price equation relating home price to distance to hazardous waste sites. They then sent a survey containing a DCE to the same home buyers, asking them to select among hypothetical houses of various prices and distances from hazardous waste sites. The authors use GMM to jointly estimate consumer preferences for distance from hazardous waste sites using as moment conditions an expression that equates marginal willingness to pay estimated from their DCE data to that estimated from the hedonic price equation. Their goals are to (1) better characterize demand for housing attributes for discrete changes away from the baseline observed in the sales data, and (2) mitigate hypothetical bias implicit in DCE data by using observed market behavior to effectively calibrate preference parameter estimates. Our motivations are similar, although our context is different.

2. Deer Hunting in Indiana

We start with background on white-tailed deer hunting in Indiana. This context is important for understanding our data and empirical approach. Hunting white-tailed deer has been a popular recreational activity in Indiana since the species' recovery from extirpation due to overhunting in the latter half of the 19th century. The Indiana legislature enacted the first game laws in Indiana

in 1857, prohibiting deer hunting from January 1 through August 1 (Michaud 1957). Since then, significant effort has been put into allocation of annual licenses by type and date of the season with both conservation and recreation value in mind.

Indiana's state game agency, the Department of Natural Resources (DNR), manages and enforces hunting regulations. The DNR categorizes hunting licenses based on hunter characteristics, reflecting age, military service status, or residency status. There are also different licenses for each season, and each season features different bag limits and equipment restrictions. Resident hunters enjoy lower license prices across all license types, while nonresident hunters pay more. Table 1 lists the attributes for the most popular license types, which we focus on here.¹

In addition to the license types mentioned previously, hunters may also purchase bonus antlerless licenses, each of which allow the harvest of one or more antlerless deer (i.e., female deer or male deer with antler spikes less than three inches long) in addition to the bag limit permitted under the other single-season licenses. Each Indiana county has different individual quotas; hunters wishing to harvest more bonus antlerless deer than the county quota allows will need to travel to a different county to do so. Indiana enforces a one antlered deer-per-season limit statewide aside from reduction zone quotas. All single-season licenses, including the first bonus antlerless license, cost \$24. Subsequent bonus antlerless licenses cost \$15 each. Finally, the DNR offers a license bundle, which allows the harvest of up to three deer (at most one antlered) across all seasons for \$65, which is less than the cost of buying each single-season license individually.

3. Survey Instrument and Data

Data for the analysis come from a mail survey of Indiana resident deer hunters, conducted following the procedures outlined by Dillman et al. (2014). A copy of the survey is available in

the Appendix. The survey comprised three parts. The first part collected information on respondents' hunting behavior, including their preferred seasons and the number of deer they want to harvest per season as well as their perceptions about disease risks in the state's deer herd. The second part comprised a DCE meant to elicit respondents' preferences towards different deer license attributes. The final part of the survey collected demographic information about the respondents.

Our DCE presented each respondent with two hypothetical licenses alongside an option not to purchase either license. Each license was given an un-descriptive label (“License A” and “License B”) and varied in the bag limit, season and equipment restrictions, and price (Table 2). The bag limit and season/equipment attribute levels were drawn from a list reflecting the current Indiana deer license structure. The price attribute levels were drawn from a range comprising 50–300% of current license prices, which correspond with deer license price levels from nearby states. We used SAS to find a *D*-optimal design ($D = 95.3$) with six blocks of ten choice sets, with each respondent being randomly assigned to one block.²

We sent two waves of surveys to each respondent two weeks apart in January 2021. In addition, each respondent received a postcard ten days after the initial mailing, reminding them to complete the survey if they had not already. Our sample comprised 7,500 resident hunters whose addresses were randomly drawn from a list of hunters who had purchased at least one deer license the previous five years. In total, 1,815 hunters (24%) returned the survey. After removing surveys with incomplete responses or presumed protest responses, the final number of usable responses was 1,398, for a final response rate of ~19%. Appendix Table A2 presents the proportion of our sample by demographic characteristic—including gender, age, and income—along with the population proportions calculated from IDNR data.

4. License Demand Estimation

We now describe our approach for estimating demand for deer hunting licenses using the data from our DCE. We start by developing a structural model of hunters' license purchasing behavior. We will use this model to derive moment conditions equating predicted license buying behavior from our model to observed, aggregate sales data that we will use in estimation.

A Model of License Choice

Assume hunter $i \in \{1, \dots, N\}$ has exogenous preferences over the number of deer (including both antlered and antlerless) that he or she wishes to harvest. In general, these preferences may depend on a vector of the hunter's personal characteristics, \mathbf{Z}_i , which includes income and has joint density $f(\mathbf{Z})$. Denote the total desired harvest as $d(\mathbf{Z}_i)$.³

For simplicity, we assume hunters make license purchasing decisions in two separate stages. In stage 1, which occurs prior to the start of the deer season, hunters face a choice set comprising the first four licenses listed in Table 1. Each permit allows hunters to hunt in distinct seasons, including archery season, firearm season, and muzzleloader season. In addition, hunters can choose to purchase a license bundle, which allows the hunter to harvest up to three deer across any season. We make a simplifying assumption that individuals that wish to hunt across multiple seasons will purchase a bundle; otherwise, they will purchase either an archery, firearm, or muzzleloader license. This implies the choice of stage 1 licenses is mutually exclusive.

Finally, we also assume the hunter can opt out, or choose no license in a given year.

In stage 2, hunters decide whether to buy bonus antlerless licenses. We assume this decision occurs conditional on (1) the hunter filling the bag limit on the license they choose in

stage 1, and (2) the desired harvest $d(\mathbf{Z}_i)$ being greater than the bag limit from their stage 1 license. We assume that the hunter buys bonus antlerless licenses sequentially—that is, they purchase each additional bonus antlerless license conditional on having filled the bag limit for a previous license purchase, and that they may purchase additional bonus antlerless licenses until their total harvest equals $d(\mathbf{Z}_i)$.

Formally, let the indirect utility from choosing license $j \in J^1 = \{\text{archery, firearm, muzzleloader, bundle}\}$ for a specific hunter i in stage 1 be $V_{ij} = U(\mathbf{X}_j^1, \mathbf{Z}_i; \boldsymbol{\theta}) + \epsilon_{ij}$, where $U(\cdot)$ is the “deterministic” portion of utility which is, in general, a function of an (observed) K -dimensional vector of license attributes \mathbf{X}_j^1 —including equipment/season restrictions, bag limits, and license prices—and \mathbf{Z}_i . The K -dimensional vector $\boldsymbol{\theta}$ includes marginal utility parameters to be estimated. The term ϵ_{ij} is a random utility shock. The deterministic component of utility from opting out, indexed as $j = 0$, comprises an alternative-specific constant (ASC) θ_0 .

The probability hunter i chooses license $j \in \{0, J^1\}$ in stage 1 is

$$\pi_j^1(\mathbf{X}^1, \mathbf{Z}_i; \boldsymbol{\theta}) = \Pr \left(U(\mathbf{X}_j^1, \mathbf{Z}_i; \boldsymbol{\theta}) + \epsilon_{ij} \geq U(\mathbf{X}_{j'}^1, \mathbf{Z}_i; \boldsymbol{\theta}) + \epsilon_{ij'} \quad \forall j' \neq j \right), \quad [1]$$

where we refer to $\mathbf{X}^1 = \{\mathbf{X}_j^1\}_{\forall j \in J^1}$ as the stage 1 “market” attribute vector to distinguish it from the DCE data we use for preference parameter estimation below. We can state demand for license j in the form of market shares using [1]. Specifically, let the probability hunter i buys license $j \in J^1$ conditional on buying a license be $\hat{\pi}_j^1(\mathbf{X}^1, \mathbf{Z}_i; \boldsymbol{\theta}) = \pi_j^1(\mathbf{X}^1, \mathbf{Z}_i; \boldsymbol{\theta}) / \sum_{j' \in J^1} \pi_{j'}^1(\mathbf{X}^1, \mathbf{Z}_i; \boldsymbol{\theta})$. The market share for license j is then

$$q_j(\mathbf{X}^1; \boldsymbol{\theta}) = \int_{Z_1} \cdots \int_{Z_K} \hat{\pi}_j^1(\mathbf{X}^1, \mathbf{Z}; \boldsymbol{\theta}) f(\mathbf{Z}) dZ_1 \cdots dZ_K. \quad [2]$$

Next, consider the demand for bonus antlerless licenses. The hunter buys these in stage 2 of their decision problem. The hunter can only buy one archery license, firearm license, muzzleloader license, or license bundle, and hence the hunter's choice set in stage 2 reduces to either opting out or buying a bonus antlerless license. The probability a hunter purchases a bonus antlerless license is

$$\pi^2(\mathbf{X}_{\text{bonus}}, \mathbf{Z}_i; \boldsymbol{\theta}) = \Pr(U(\mathbf{X}_{\text{bonus}}, \mathbf{Z}_i; \boldsymbol{\theta}) + \epsilon_{i,\text{bonus}} \geq \theta_0 + \epsilon_{i0}). \quad [3]$$

We assume for simplicity that this probability is independent of the number of deer already harvested. Recall, however, that the price of the first bonus antlerless license is \$24, whereas each subsequent license costs \$15. We therefore use $\mathbf{X}_{\text{bonus},1}$ to refer to the attributes of the first bonus license and $\mathbf{X}_{\text{bonus},2}$ to refer to the attributes of all subsequent licenses.

Denote the bag limit for stage 1 license j as \bar{d}_j . We assume the probability a hunter harvests a deer (antlered or antlerless) is constant, independent across harvests, and equal to p . We infer a value of $p = 0.45$ from DNR harvest data. Hence, the probability that a hunter fills their bag limit for license $j \in J^1$ is $p^{\bar{d}_j}$. If $d(\mathbf{Z}_i) > \bar{d}_j$ such that the hunter wants to harvest more deer than allowed under their stage 1 license choice, the hunter will have the opportunity to purchase one or more bonus antlerless licenses. The expected number of bonus antlerless licenses purchased per hunter is

$$q_{\text{bonus}}(\mathbf{X}^2; \boldsymbol{\theta}) = \int_{Z_1} \cdots \int_{Z_K} \sum_{j \in J^1} \sum_{\delta} \delta \Pr(\delta | d(\mathbf{Z}) - \bar{d}_j) \pi_j^1(\mathbf{X}^1, \mathbf{Z}, \boldsymbol{\theta}) p^{\bar{d}_j} f(\mathbf{Z}) dZ_1 \cdots dZ_K, \quad [4]$$

where

$$\Pr(\delta | \bar{d}_j - d(\mathbf{Z})) = \begin{cases} \pi^2(\mathbf{X}_{\text{bonus},1}, \mathbf{Z}; \boldsymbol{\theta}) [\pi^2(\mathbf{X}_{\text{bonus},2}, \mathbf{Z}; \boldsymbol{\theta}) p]^{\delta-1} [1 - p\pi^2(\mathbf{X}_{\text{bonus},2}, \mathbf{Z}; \boldsymbol{\theta})] & 0 < \delta < d(\mathbf{Z}) - \bar{d}_j \\ \pi^2(\mathbf{X}_{\text{bonus},1}, \mathbf{Z}; \boldsymbol{\theta}) [\pi^2(\mathbf{X}_{\text{bonus},2}, \mathbf{Z}; \boldsymbol{\theta}) p]^{\delta-1} & \delta = d(\mathbf{Z}) - \bar{d}_j \\ 0 & \text{otherwise.} \end{cases}$$

(See Appendix for derivation.)

Preference Parameter Estimation

Equations [2] and [4] constitute the demand system for deer licenses. We can estimate these equations in a few ways. The standard approach, given our choice data, is to specify a distribution for the error terms ϵ_{ij} —say, type-1 extreme value, so that the choice probability in [1] takes the standard conditional logit form $\pi_j(\mathbf{X}_{it}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta}) = e^{U(\mathbf{X}_{ijt}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta})} / \sum_{j'} e^{U(\mathbf{X}_{ij't}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta})}$. Note that the attribute data $\mathbf{X}_{it}^{\text{CE}} = \{\mathbf{X}_{ijt}^{\text{CE}}\}_{\forall j}$ come from our optimal experimental design—these are the license attributes survey respondent i was shown in choice occasion t —and hence are distinct from the real-world, market license attributes, \mathbf{X}^1 and \mathbf{X}^2 , described in the previous section. We can then estimate the marginal utility parameters $\boldsymbol{\theta}$ via maximum likelihood (MLE). Formally, we specify the log-likelihood as $\mathcal{L} = \sum_{i \times t \times j} \psi_{itj} \ln \pi_j(\mathbf{X}_{it}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta})$, where $\psi_{itj} = 1$ if survey respondent i chose alternative j in choice occasion t and zero otherwise. The estimated parameters, which we denote $\hat{\boldsymbol{\theta}}_{\text{MLE}}$, solve the first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} = \sum_{i \times t \times j} \psi_{itj} \left[\frac{\partial U(\mathbf{x}_{ijt}^{\text{CE}}, \mathbf{z}_i; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - \sum_{j'} \frac{\partial U(\mathbf{x}_{ij't}^{\text{CE}}, \mathbf{z}_i; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \pi_{j'}(\mathbf{X}_{it}^{\text{CE}}, \mathbf{z}_i; \hat{\boldsymbol{\theta}}_{\text{MLE}}) \right] = 0. \quad [5]$$

We can then predict demand by calculating [2] and [4] using $\hat{\boldsymbol{\theta}}_{\text{MLE}}$, attributes of the market licenses \mathbf{X}^1 and \mathbf{X}^2 , and knowledge of the population distribution of \mathbf{Z}_i . (Estimates of a limited set of hunter characteristics are frequently available from state or federal-level wildlife management agencies.) Note, however, that the resulting demand estimates will match observed demand for the existing suite of licenses only under very restrictive conditions. Specifically, suppose we showed a representative sample of respondents a single choice set with the entire suite of available, real-world licenses, each of which had levels set at their real-world values. If the data from this exercise, which we'll call \mathbf{X}_i^{CE} , comprised a full set of ASCs such that $X_{ijl}^{\text{CE}} = 1$ for $j = l$ and zero otherwise, and utility is linear in these ASCs, then [5] would reduce to

$$\begin{aligned} \sum_i [\psi_{il} - \pi_l(\mathbf{X}_i^{\text{CE}}, \mathbf{z}_i; \hat{\boldsymbol{\theta}}_{\text{MLE}})] &= 0 \\ \Rightarrow \frac{\sum_i \psi_{il}}{N} &= \frac{\sum_i \pi_l(\mathbf{X}_i^{\text{CE}}, \mathbf{z}_i; \hat{\boldsymbol{\theta}}_{\text{MLE}})}{N} \quad \forall l, \end{aligned} \quad [5']$$

implying that the mean estimated probability of choosing alternative l (the right-hand side of the second equation in [5']) exactly equals the observed aggregate choice share (the left side). This feature of conditional logit models is well-known (Klaiber and von Haefen 2019). However, our data has several characteristics that are common among those collected from DCEs that would prevent our estimated choice shares from matching observed ones. First, we showed our respondents choice sets comprising two alternative licenses and a “no choice” option, rather than the full suite of licenses. These alternatives were “generic” in the sense described by Holmes and Adamowicz (2003): they were given undescriptive labels (“License A” and “License B”) and all

attributes of each alternative varied between choice occasions. We made these choices for statistical efficiency, and indeed this basic format is common among DCEs. As a consequence, though, there is no meaningful variation in alternatives upon which to identify any ASCs. Even if respondents were shown labeled choice sets such that we could estimate ASCs, we would not expect the estimated shares from our choice experiment to match real-world shares. To see why, consider a simplified case in which a respondent who does not own a bow is shown a choice set containing two archery licenses and an opt-out option. It is likely this respondent would choose the opt-out option, whereas in a real-world setting this person would simply choose a different, non-archery license. As a result, the choice experiment would overestimate this respondent's propensity to choose the opt-out option.⁴ Finally, our sample is not perfectly representative of the population; in particular, the sample is older and has fewer middle-income individuals relative to the population.⁵ Even if we were to adjust for this (e.g., by applying appropriate post-stratification weights in estimation), we would not expect our model to replicate observed market shares for the reasons cited previously.

We can force our model to replicate observed market shares by calibrating utility to observed market shares, which we denote q_j^* , $j \in J^1$. The standard approach to calibration is to (1) add an ASC α_j to the estimated utility functions, such that utility from alternative j is $\alpha_j + U(\mathbf{X}_j^1, \mathbf{Z}_i; \boldsymbol{\theta}) + \epsilon_{ij}$; (2) use these augmented utilities to calculate market shares $q_j(\mathbf{X}^1; [\boldsymbol{\alpha} \hat{\boldsymbol{\theta}}_{\text{MLE}}])$ and $q_{\text{bonus}}(\mathbf{X}^2; [\boldsymbol{\alpha} \hat{\boldsymbol{\theta}}_{\text{MLE}}])$ from [2] and [4]; and (3) solve for the values of these constants that replicate observed demand:

$$\begin{aligned} \mathbf{q}(\mathbf{X}^1; [\boldsymbol{\alpha} \hat{\boldsymbol{\theta}}_{\text{MLE}}]) - \mathbf{q}^* &= \mathbf{0} \\ q_{\text{bonus}}(\mathbf{X}^2; [\boldsymbol{\alpha} \hat{\boldsymbol{\theta}}_{\text{MLE}}]) - q_{\text{bonus}}^* &= 0 \end{aligned} \tag{6}$$

(e.g., Naidoo and Adamowicz 2005). We chose not to apply this calibration approach to our problem. This is for a few reasons. First, the interpretation of the α_j terms is not straightforward in our context. In standard choice settings, ASCs capture preferences over attributes specific to a particular alternative that are not included in the model specification. However, hunting licenses are completely defined by price, bag limits and season/equipment restrictions; it is not clear what other attributes may be captured by an ASC, and hence including them in the model *ex post* for the purpose of calibration is hard to defend. Conversely, even if the model was mis-specified, estimating a model and then calculating a term outside the estimation to force replication seems a bit *ad hoc*. Justifying the term's presence by stating that it can compensate for possible model misspecification just leads to further questions about the quality of the original estimates. Finally, this calibration procedure cannot be used when simulating demand for entirely new license types since there is no existing data with which to calibrate utility.

Instead, we use a GMM-based approach following Imbens and Lancaster (1994) that uses the observed market shares to inform estimation of the preference parameters. Formally, we use equations [2] and [4] to specify a system of $K + 4$ moment conditions

$$\begin{aligned}
 \mathbf{g}_1(\mathbf{X}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta}) &= \frac{1}{I \times T} \sum_{i \times t \times j} \psi_{itj} \left[\frac{\partial U(\mathbf{x}_{ijt}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} - \sum_{j'} \frac{\partial U(\mathbf{x}_{ij't}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \pi_{j'}(\mathbf{X}_{it}^{\text{CE}}, \mathbf{Z}_i; \boldsymbol{\theta}) \right] = \mathbf{0} \\
 \mathbf{g}_2(\mathbf{X}^1; \boldsymbol{\theta}) &= \mathbf{q}(\mathbf{X}^1; \boldsymbol{\theta}) - \mathbf{q}^* = \mathbf{0} \\
 g_3(\mathbf{X}^2; \boldsymbol{\theta}) &= q_{\text{bonus}}(\mathbf{X}^2; \boldsymbol{\theta}) - q_{\text{bonus}}^* = 0,
 \end{aligned} \tag{7}$$

then choose the marginal utility parameters $\hat{\boldsymbol{\theta}}_{\text{GMM}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \mathbf{G}'\mathbf{W}(\boldsymbol{\theta})\mathbf{G}$, where $\mathbf{G} =$

$[\mathbf{g}_1(\cdot) \mathbf{g}_2(\cdot) g_3(\cdot)]'$ and $\mathbf{W}(\boldsymbol{\theta})$ is a $K \times K$ weight matrix. Note that including $\mathbf{g}_2(\cdot)$ in system [7]

serves a similar role as estimating ASCs would if the latter were actually estimable. In contrast to a model with ASCs, replication under the GMM approach will not be exact, but we show later that our predicted shares are still quite close to observed shares.

It is worth noting how parameter identification differs between the standard estimation approach and the GMM-based approach we use here. Under the GMM approach, the market moment conditions $\mathbf{g}_2(\cdot)$ and $g_3(\cdot)$ pin down the relative magnitudes of the total utility from each combination of attributes observed in the market. The DCE data identifies differences in the relative magnitudes of the contributions to total utility from each individual license attribute. Under the standard approach, the DCE data performs both tasks. It is the former task that allows the GMM approach to more closely replicate the observed demand; because subjects in the DCEs never see the entire suite of actual licenses available in the same choice set, the standard approach cannot identify differences in the magnitudes of total utility across license types as effectively as the GMM approach.

We must specify the weight matrix $\mathbf{W}(\cdot)$ before proceeding with estimation. We set $\mathbf{W}(\cdot) = \mathbf{I}_K$, where \mathbf{I}_K is a K -dimensional identity matrix. Doing so results in estimates $\hat{\boldsymbol{\theta}}_{\text{GMM}}$ that are consistent, if not asymptotically efficient.⁶

5. Results

We estimate marginal utility parameters $\boldsymbol{\theta}$ assuming the indirect utility from license j depends only on the attributes of each license and not on the hunters' personal characteristics \mathbf{Z}_i .

Formally, we specify utility for license j as $U_j(\mathbf{X}_j; \boldsymbol{\theta}) = \mathbf{X}_j \boldsymbol{\theta}$, where the \mathbf{X}_j are dummy-coded variables describing the bag limits and equipment/season restrictions for each license (see Table 17) and price.⁸ The first and third columns of Table 3 show $\hat{\boldsymbol{\theta}}_{\text{MLE}}$ and $\hat{\boldsymbol{\theta}}_{\text{GMM}}$, respectively. Figure

1 shows estimated and actual market shares for the actual suite of licenses offered to hunters, along with 95% confidence intervals.

[Insert Table 1 about here]

[Insert Figure 1 about here]

The second and fourth columns of Table 3 show the standard errors for each estimate, clustered at the individual level. All parameter estimates are significantly different from zero except for the 1 antlered + 1 antlerless bag limit level in both models; this simply means that respondents receive no extra utility from being able to harvest a second antlerless deer, all else equal. Likewise, all estimates have the same signs. The signs are also sensible; hunters gain the greatest utility from the ability to harvest one antlered and two antlerless deer (the base category for the bag limit dummy variables), with smaller bag limits generating relatively less utility. Likewise, the ability to hunt across any season (the base case for the equipment/season dummy variables) generates more utility than being restricted to hunt within a single season—whether archery, muzzleloader, or firearm season. The parameter on license price is almost identical across models. However, the estimates of other license attributes differ in some key ways. Notably, the parameter estimates for the one antlered, one antlerless bag limit dummy and the archery equipment dummies are considerably smaller under the GMM approach. These attributes correspond to the real-world archery season. Figure 1 shows that the MLE model badly overestimates market shares for this license. In contrast, the GMM parameter estimates assign much less utility to these attributes, resulting in a closer match between predicted and observed

shares. Similarly, the GMM model also generates a closer match between predicted and observed shares for the license bundle.

The model parameters are also more precisely estimated under the GMM approach. The standard errors of our GMM estimates as a fraction of the estimated parameters are much smaller than under the MLE approach. The relative efficiency of the GMM approach is well-known; indeed, it was the main motivation behind Imbens and Lancaster's (1994) original paper.

We highlight the advantage of the GMM approach with a counterfactual analysis in which we simulate changes to deer license demand following changes to license offerings. Specifically, we simulate the effect of (1) changing bag limits for the muzzleloader and archery licenses to a single antlered deer, and (2) increasing single-season license prices from \$24 to \$40–\$60 each and license bundles from \$65 to \$90, with no discount for bonus antlerless licenses. Officials with the DNR are considering this scenario as a means of increasing agency revenues and improving their ability to manage antlered and antlerless deer populations within the state. For each combination of license prices, we calculate compensating variation (CV), predicted changes to DNR revenues, and the predicted change in demand for each license type relative to baseline license demand under the MLE and GMM models. For the MLE model, we calibrate baseline demand for each existing license type as in [6] so as to not unfairly stack the deck against this approach. Details about the simulation are in the Appendix; here, we only report the simulation results and provide the basic intuition behind our findings.

Figure 2 shows the resulting change in demand for each of the single license types (archery, firearm, muzzleloader, and bonus antlerless) assuming the price of a bundle is fixed at \$90. The main effect of this change in license structure is to make archery licenses less appealing; note from Table 3 that the marginal utility from the one antlered deer bag limit is

considerably smaller in magnitude than that from the one antlered and one antlerless bag limit. The effect is to cause hunters to substitute away from archery licenses. The response to the change in structure is similar under each model, although the MLE model predicts a larger number of hunters will stop buying archery licenses.

[Insert Figure 2 about here]

Figure 3a shows the CV for the simulated license changes per person under the GMM and MLE models. It is negative, implying hunters are worse-off following the change to license structure. The GMM estimates are larger (i.e., less negative) than the MLE estimates on average, although the differences in means are small and, for the most part, not statistically significantly different.

[Insert Figure 3 about here]

Figure 3b shows the simulated change in IDNR license revenues following the change in license offerings. Agency revenues increase with the single license price, even though overall demand for licenses decreases. This implies license demand is inelastic and suggests some scope for increasing license prices from their baseline levels. The predicted increase in revenue is smaller under the GMM model due to more hunters opting out, although the differences are not statistically significant.

Notably, the confidence intervals around the CV and revenue changes are much smaller for the GMM model. This is despite the fact that we calibrated the MLE model to replicate

baseline demand. The explanation for this result is straightforward. We calculated the confidence intervals in Figure 3 using the Krinsky-Robb procedure by taking random draws of the model parameters from their estimated distributions under the MLE and GMM models. Denote these parameters as $\widehat{\boldsymbol{\theta}}_{\text{MLE}}^r$ and $\widehat{\boldsymbol{\theta}}_{\text{GMM}}^r$, where the superscript r denotes the draw. Given the distributional assumption on ϵ_{ij} , we used these parameter draws to calculate expected baseline surplus under the MLE and GMM models, respectively:

$$S_{\text{MLE}}^{0,r} = \frac{1}{|\theta_{\text{price}}^r|} \left[\ln \left(\sum_{j \in J^1} e^{\alpha_j^r + \mathbf{X}_j^1 \widehat{\boldsymbol{\theta}}_{\text{MLE}}^r} \right) + q_{\text{bonus}}(\mathbf{X}^2; [\boldsymbol{\alpha}^r \widehat{\boldsymbol{\theta}}_{\text{MLE}}^r]) \ln \left(\sum_{j \in J^2} e^{\alpha_j^r + \mathbf{X}_j^2 \widehat{\boldsymbol{\theta}}_{\text{MLE}}^r} \right) \right], \quad [8a]$$

$$S_{\text{GMM}}^{0,r} = \frac{1}{|\theta_{\text{price}}^r|} \left[\ln \left(\sum_{j \in J^1} e^{\mathbf{X}_j^1 \widehat{\boldsymbol{\theta}}_{\text{GMM}}^r} \right) + q_{\text{bonus}}(\mathbf{X}^2; \widehat{\boldsymbol{\theta}}_{\text{GMM}}^r) \ln \left(\sum_{j \in J^2} e^{\mathbf{X}_j^2 \widehat{\boldsymbol{\theta}}_{\text{GMM}}^r} \right) \right], \quad [8b]$$

where θ_{price} is the parameter on price (such that $|\theta_{\text{price}}|$ is the marginal utility of income) and J^2 is the hunter's choice set of licenses available in the second-round, comprising only the opt-out and bonus antlerless licenses. Note that the only difference between these two surplus measures (other than the parameter values themselves) is the presence of the α_j^r terms in [8a]. These terms are calculated from the calibration procedure in the MLE model, but not the GMM model. Next, let $\check{\mathbf{X}}^1$ and $\check{\mathbf{X}}^2$ be the first- and second-stage license attributes following the simulated change in bag limits and price. Surplus following this change, $S_{\text{MLE}}^{1,r}$ and $S_{\text{GMM}}^{1,r}$, is given by [8] after substituting $\check{\mathbf{X}}^1$ and $\check{\mathbf{X}}^2$ for the market attributes \mathbf{X}^1 and \mathbf{X}^2 . CV is then

$$S_k^{1,r} - S_k^{0,r}, k \in \{\text{MLE}, \text{GMM}\}. \quad [9]$$

The baseline surplus $S_k^{0,r}$ will be similar under both the GMM and MLE models. This is because the calibration procedure used in the MLE model will result in utility functions $\alpha_j^r + \mathbf{X}_j^1 \hat{\boldsymbol{\theta}}_{\text{MLE}}^r$ and $\alpha_j^r + \mathbf{X}_j^2 \hat{\boldsymbol{\theta}}_{\text{MLE}}^r$ that replicate the baseline, observed behavior. Baseline surplus—which depends on these utility functions—would therefore not vary, regardless of the parameter draws. Of course, we do not calibrate the GMM model estimates. However, the estimated variance of the GMM estimates is relatively small, and since the GMM estimates reliably replicate observed behavior, we would expect the distribution of baseline surplus to be fairly tight in the GMM model as well. Given that the calibration procedure (in the MLE model) and the relatively precise estimates of $\boldsymbol{\theta}$ (in the GMM model) effectively pin down the second term in [9], the variation in CV must arise primarily from variation in the first term of [9]. To the extent that the variance of the MLE estimates is larger than for the GMM estimates, we would expect the variance of this surplus term to be greater in the MLE model as well.

The explanation for the wider confidence intervals on the revenue calculations follows a similar logic. Taken together, these results suggest that the GMM approach may provide more precise predictions than calibration-based procedures. Furthermore, the GMM approach is more broadly applicable; the calibration procedure cannot be reliably applied to simulate counterfactual scenarios examining demand for novel license types given the lack of market data to calibrate the model.

6. Conclusion

Estimating demand for recreational licenses and their attributes is complicated by the fact that there is typically no meaningful variation in license prices or structure that one can exploit.

DCEs offer a means of experimentally varying license attributes to identify preference

parameters. An issue with choice models estimated from DCEs is that the predicted market shares are unlikely to replicate observed market shares, reducing the credibility of counterfactual simulations performed with these models. It is possible to calibrate the estimated choice models. A common approach is to add ASCs to the estimated choice model *ex post*, then solve for the value of these ASCs that replicate observed aggregate choice data. Calibration is valid when simulating demand for complex, multi-attribute goods that are similar to the existing goods for which one has aggregate data (net of changes to the observed attributes from the choice experiment). Calibration cannot be used in cases like the one we examine here, which involves simulating demand for novel goods for which observed aggregate data does not exist. We use a GMM-based approach to estimate a choice model using a combination of DCE data and observed license sales data. Our estimated model much more closely replicates observed market share data relative to a standard conditional logit model, adding credibility when performing counterfactual analyses of license structure changes. This approach can also be used to estimate demand for novel goods—here, new types of hunting licenses—for which no aggregate choice data exists.

Our approach is slightly more data-intensive than the standard approach to estimating DCE choice models. However, this is of little consequence as resource management agencies should track license sales data; indeed, this information is routinely reported to sportsmen and – women in publicly-available agency publications.

Acknowledgements

This project was funded by Wildlife Restoration Grant F20AF10970,W-48-R-04, in cooperation with the Indiana Department of Natural Resources, Division of Fish and Wildlife, and the U.S. Fish and Wildlife Service Wildlife Restoration Grant Program.

References

- Caudell, J. N. and O. D. L. Vaught. 2019. "Indiana White-Tailed Deer Report." Department of Natural Resources, Bloomington, Indiana.
- Dillman, D. A., J. D. Smyth, and L. M. Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Fourth ed. Wiley: Hoboken, New Jersey.
- Erickson, D. C. Reeling, and J. G. Lee. 2019. "The Effect of Chronic Wasting Disease on Resident Deer Hunting License Demand in Wisconsin." *Animals* 9(12):1096.
- Greene, W.H. 2012. *Econometric Analysis*. 7 ed. Pearson: Boston.
- Hausman, J. and D. McFadden. 1984. "Specification Tests for the Multinomial Logit Model." *Econometrica* 52(5):1219–1240.
- Holmes, T.P. and W. L Adamowicz. 2003. "Attribute-Based Methods." In Champ, P.A., K.J. Boyle, and T.C. Brown (Eds.) *A Primer on Nonmarket Valuation*. Kluwer: Dordrecht.
- Imbens, G. W., and T. Lancaster. 1994. "Combining Micro and Macro Data in Microeconomic Models." *The Review of Economic Studies* 61(4): 655–680.
- Klaiber, H.A. and R.H. von Haefen. 2019. "Do Random Coefficients and Alternative Specific Constants Improve Policy Analysis? An Empirical Investigation of Model Fit and Prediction." *Environmental and Resource Economics* 73:75–91.

- Loomis, J., C. Pierce, and M. Manfredo. 2000. "Using the Demand for Hunting Licenses to Evaluation Contingent Valuation Estimates of Willingness to Pay." *Applied Economics Letters* 7(7):435–438.
- Louviere, J.J., D.A. Hensher, and J.D. Swait. 2000. *Stated Choice Methods: Analysis and Application*. Cambridge University Press.
- Mackenzie, J. 1990. "Conjoint Analysis of Deer Hunting." *Northeastern Journal of Agricultural and Resource Economics* 19(2):109–117.
- Michaud, H.H. 1957. "History of the Early Development of Game Regulations in Indiana." *Proceedings of the Indiana Academy of Science* 67:256–259.
- Mingie, J.C., N.C. Poudyal, J.M. Bowker, M.T. Mengak, and J.P. Siry. 2017. "Big Game Hunter Preferences for Hunting Club Attributes: A Choice Experiment." *Forest Policy and Economics* 78:98–106.
- Naidoo, R. and W. Adamowicz. 2005. "Biodiversity and Nature-Based Tourism at Forest Reserves in Uganda." *Environment and Development Economics* 10(2):159–178.
- Phaneuf, D.J., L.O. Taylor, and J.B. Braden. 2013. "Combining Revealed and Stated Preference Data to Estimate Preferences for Residential Amenities: A GMM Approach." *Land Economics* 89(1):30–52.
- SAS Institute, Inc. 2014. *SAS/QC[®] 13.2 User's Guide*. SAS Institute, Inc.: Cary, NC
- Serenari, C., J. Shaw, R. Myers, and D.T. Cobb. 2019. "Explaining Deer Hunter Preferences for Regulatory Changes using Choice Experiments." *The Journal of Wildlife Management* 83(2):446–456.

- Schroeder, S.A., D.C. Fulton, L. Cornicelli, and S.S. Merchant. 2018. “Discrete Choice Modeling of Season Choice for Minnesota Turkey Hunters.” *The Journal of Wildlife Management* 82(2):457–465.
- Train, K.E. 2009. *Discrete Choice Methods with Simulation* (2nd ed.). Cambridge University Press.
- Vossler, C.A., M. Doyon, and D. Rondeau. 2012. “Truth in Consequentiality: Theory and Field Evidence on Discrete Choice Experiments.” *American Economic Journal: Microeconomics* 4(4):145–171.
- Whitehead, J.C., S.K. Pattanayak, G.L. Van Houtven, and B.R. Gelso. 2008. “Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Ecological Services: An Assessment of the State of the Science.” *Journal of Economic Surveys* 22(5):872–908.

Table 1

Market License Attributes

Season/equipment restrictions	Valid dates (2021 season)	Price	Bag limit
	Oct. 1–Jan. 2		1 antlered, 1 antlerless OR 2 antlerless
Archery		\$24	
Firearm	Nov. 13–28	\$24	1 antlered
Muzzleloader	Dec. 4–19	\$24	1 antlered
Bundle	Season long	\$65	1 antlered + 2 antlerless
Bonus antlerless	Season long (except in certain counties)	24 for first one, \$15 thereafter	1 antlerless (up to county quota)

Table 2

License Attributes and Levels used in the Choice Experiment

Attributes	Levels
Price (\$)	12, 19.20, 31.20, 36, 72
Bag limit	1 antlered, 1 antlerless, 1 antlered + 1 antlerless or 2 antlerless, 1 antlered + 2 antlerless
Season/equipment restrictions	Archery, muzzleloader, firearm, any

Table 3

Parameter Estimates for Indiana Resident Deer License Demand Models

Variable	(1)		(2)	
	MLE		GMM	
	Estimate	SE	Estimate	SE
Price	-0.039	0.006	-0.039	0.006
Opt-out	-3.632	0.737	-3.994	0.459
Bag limit				
One antlered	-1.516	0.345	-1.497	0.304
One antlerless	-2.183	0.506	-2.213	0.289
One antlered, one antlerless	-0.294	0.227	-0.654	0.391
One antlered, two antlerless (base case)	—	—	—	—
Equipment/season				
Archery	-1.508	0.316	-1.948	0.167
Muzzleloader	-2.151	0.408	-2.710	0.106
Firearm	-0.732	0.273	-0.962	0.072
Any (base case)	—	—	—	—

Note: Standard errors are clustered at the individual level. A total of 1,398 unique individuals and 13,752 choice occasions are in our data set.

Figure 1

Estimated and observed market shares of Indiana deer licenses

Figure 2

Mean change in demands for different license types under MLE and GMM models following change in bag limit to one antlered deer for archery, muzzleloader licenses

Figure 3

Changes to a) agency revenue and b) compensating variation following change in bag limit to one antlered deer for archery, muzzleloader licenses

¹ Archery, firearm, muzzleloader and bundle licenses comprised >90% of adult resident license sales in 2019 (Caudell and Vaught 2019).

² Specifically, we used the OPTEX algorithm in SAS to identify a D -optimal design. This algorithm takes as inputs the number of choice sets and blocks we wish to include in the design, and then chooses the one that maximizes the determinant of the information matrix $|\mathbf{X}'\mathbf{X}|$ —where \mathbf{X} denotes the design attributes—which is proportional to the inverse of the covariance matrix in linear models. Other approaches for designing experiments exist. In particular, D -efficient designs tend to be more efficient but require prior knowledge of the parameters to be estimated in order to generate efficient designs. For this reason, D -optimality is commonly used as a design criterion even for nonlinear models like ours (SAS Institute, Inc. 2014). We chose the number of choice sets and blocks to make the D -optimality as close to 100 as possible while keeping the experiment to a reasonable size.

³ In general, the desired harvest $d_i = d(\mathbf{Z}_i)$ could depend on the hunter's characteristics, including age, income, education level, and hunter experience level. While we have survey data on these characteristics, we find that none of them influence d_i and hence treat this as a scalar in our analysis.

⁴ As an aside, one may wonder why we did not just estimate a choice model using actual license sales data from the DNR, as doing so (including a full set of ASCs as described above) would obviate the need for our more complicated approach. Estimating such a model would not be feasible in our case because

prices are perfectly correlated to license type. Hence, the ASCs would not be identified due to collinearity. This is a common feature of regulated goods like recreational permits and not specific to our research setting. Practically speaking, this type of data may also be restricted for privacy or other reasons.

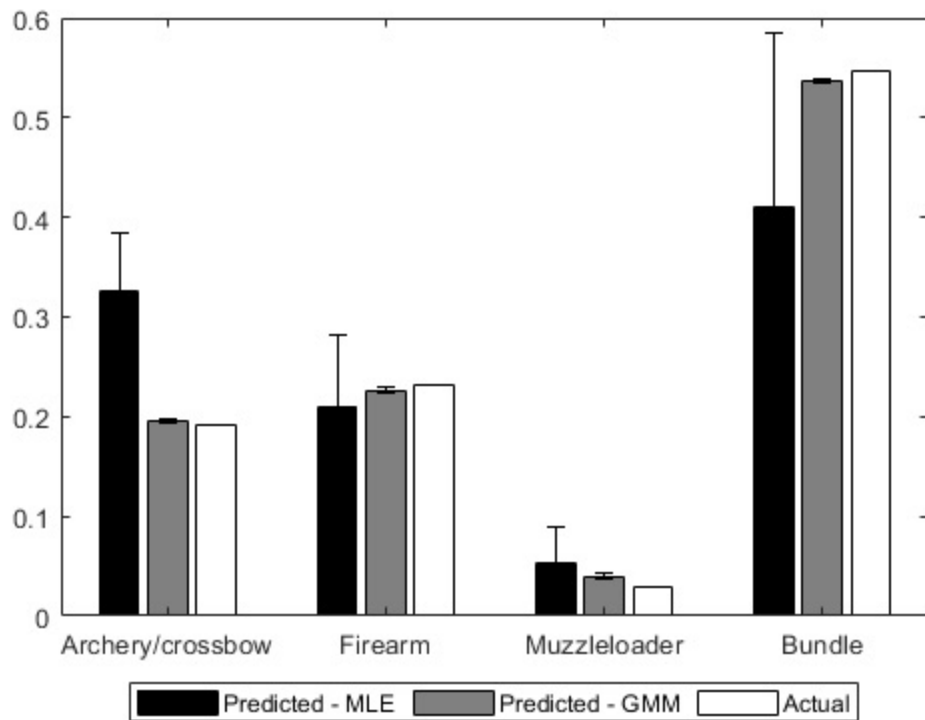
⁵ The IDNR does not collect income data for its hunters. We calculated population income shares using census block group-level median household income from the U.S. Census Bureau’s 2019 American Community Survey (see Appendix Table A2, note a). For this reason, one should be cautious when comparing the (calculated) population and (reported) sample income shares.

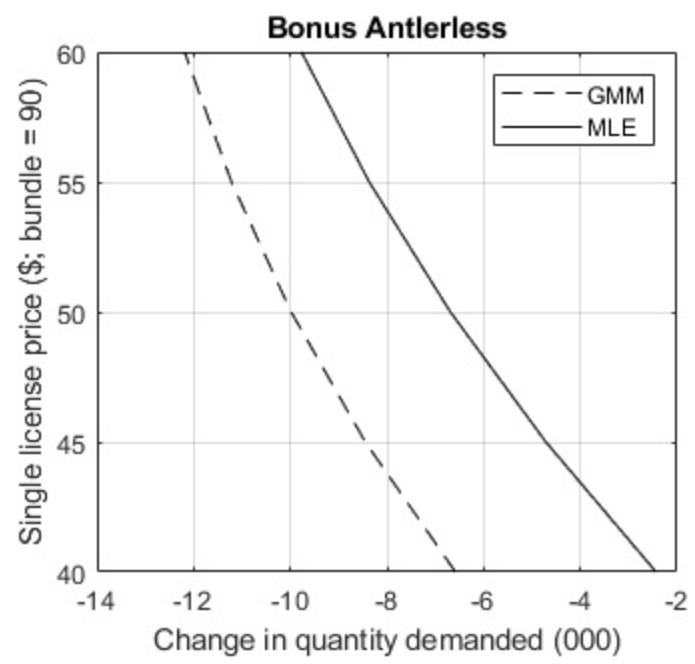
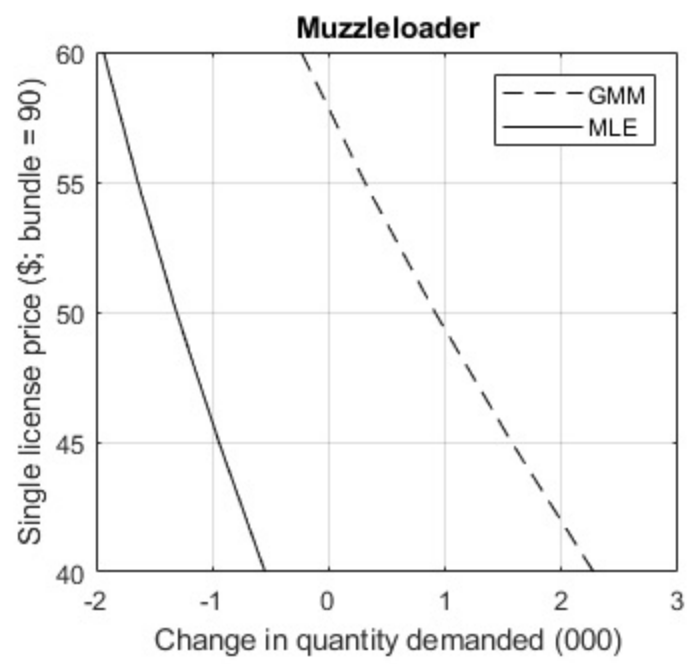
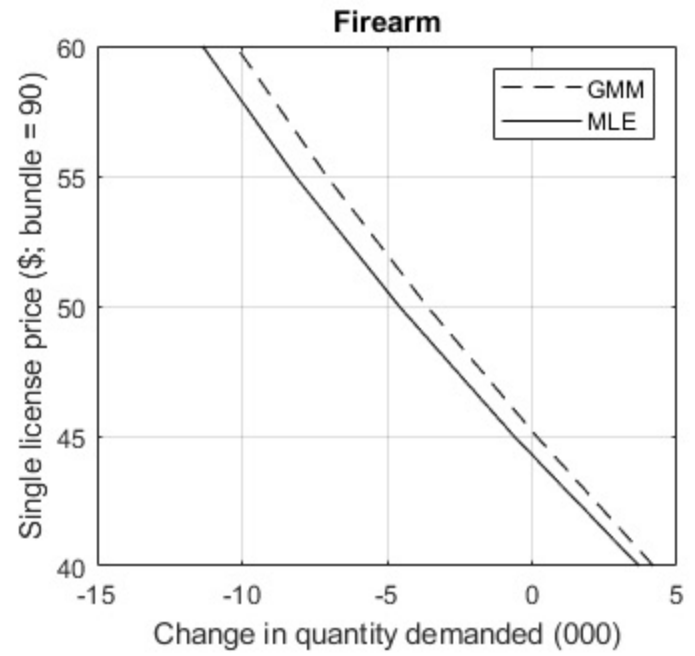
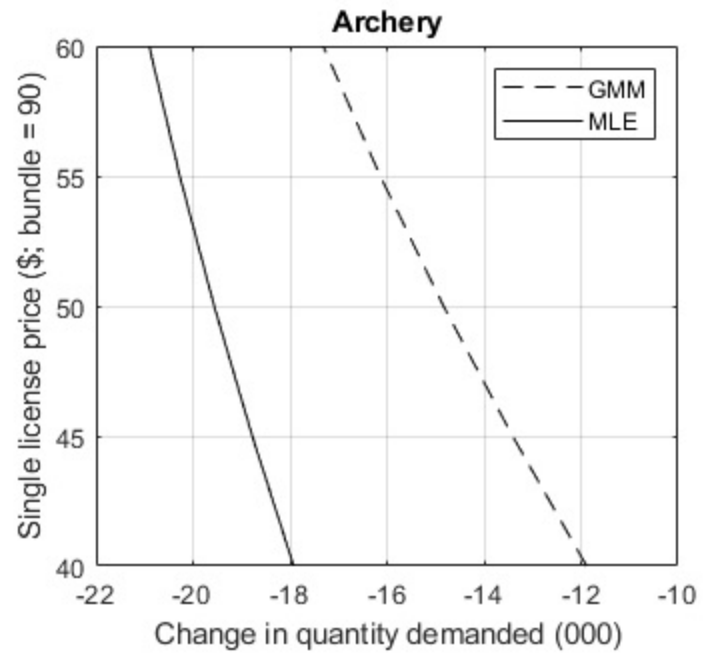
⁶ It is well known that the asymptotically efficient weight matrix equals the inverse of the asymptotic covariance of the moment vectors (Greene 2012). The market moment conditions $\mathbf{g}_2(\cdot)$ and $\mathbf{g}_3(\cdot)$ are effectively derived from different data than $\mathbf{g}_1(\cdot)$, and as a result we set the covariance terms between these groups of moments to 0. This leads to numerical instabilities when calculating the inverse of the resulting covariance matrix. Hence, we chose to use $\mathbf{W}(\cdot) = \mathbf{I}_K$.

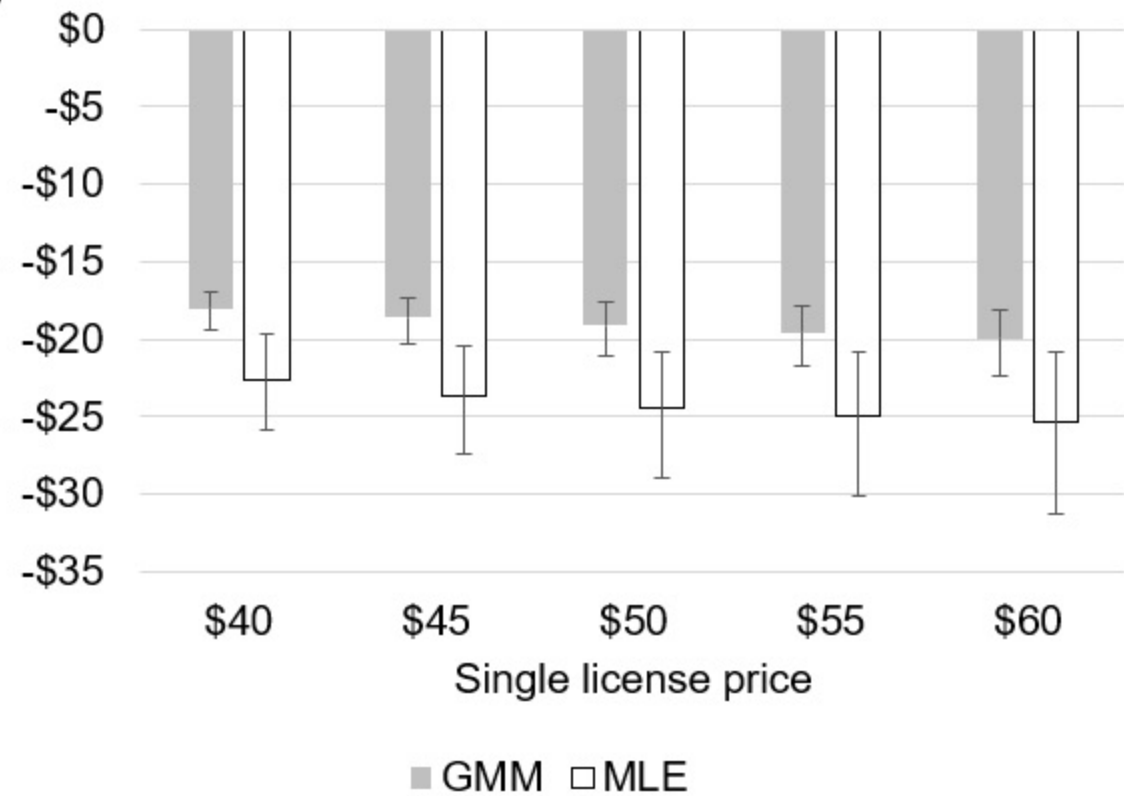
⁷ The real-world bag limits for the muzzleloader license and the license bundle are slightly different from those shown in Table 1: the muzzleloader license allows hunters to harvest one species of either sex, whereas the license bundle allows harvest of one antlered and two antlerless deer or three antlerless deer. We omitted these as levels from our choice experiment to reduce the size of our experimental design and because the IDNR is considering eliminating bag limits that allow harvest of either sex deer—hence, these bag limits would not be useful in our counterfactual analyses. We set the muzzleloader bag limit to be one antlered deer. This reflects the fact that two-thirds of the muzzleloader license buyers in our survey data only want to harvest a single, antlered deer, while the remaining buyers want to harvest one antlered *and* one antlerless deer (e.g., by buying a bonus antlerless license). We set the bundle bag limit to be one antlered and two antlerless deer for the same reason; less than one percent of bundle license buyers in our data do not want to harvest an antlered deer, implying the intended use of the bundle is to harvest one antlered deer and up to two antlerless deer.

⁸ This specification implies that our choice model takes the standard conditional logit form. Conditional logit models assume preference homogeneity and are susceptible to the independence of irrelevant alternatives (IIA), which implies unrealistic substitution patterns in the estimated model. Hausman and McFadden (1984) derive a specification test to determine whether the conditional logit model is appropriate given the choice data. We conducted this test in Appendix Section A.2 and find that the IIA assumption under the conditional logit model does not hold in our data. This finding suggests generalizing our modeling approach to a latent class or mixed logit model, neither of which suffer from the IIA assumption. These models also allow for preference heterogeneity among respondents. Incorporating either of these models into our approach would be straightforward; we would simply need to replace $\mathbf{g}_1(\cdot)$ in [7] with the latent class or mixed logit gradients. However, we opted not to do this. We

feel the more useful comparison of model prediction is between the GMM-based approach and a conditional logit model. As we explained previously, it is possible to get perfect in-sample prediction from a conditional logit model, but not from a latent class or mixed logit model. Hence, explaining the merits of the GMM approach relative to a conditional logit model makes for the most honest comparison of possible approaches.





**b)**