

# Analyzing Data from Social Dilemma Experiments: A Bayesian Comparison of Parametric Estimators

**Klaus Moeltner\***

Department of Resource Economics,  
University of Nevada, Reno

**James J. Murphy**

Department of Economics,  
University of Alaska Anchorage

**John K. Stranlund**

Department of Resource Economics,  
University of Massachusetts-Amherst

**Maria Alejandra Velez**

Facultad de Administración  
Universidad de Los Andes, Bogotá, Colombia

*First Submission, October 14, 2008*

\* Contact Information:

Klaus Moeltner

Department of Resource Economics / MS 204

University of Nevada, Reno

Reno, NV 89557-0105

phone: (775) 784-4803

fax: (775) 784-1342

e-mail: moeltner@cabnr.unr.edu

We thank Maria Claudia Lopez and members of the Faculty of Environmental and Rural Studies at Javeriana University in Bogotá, Colombia for providing outstanding support for our field experiments. We are also indebted to WWF-Colombia for coordinating the fieldwork in the Pacific Region. We received valuable comments and suggestions from Juan Camilo Cardenas, Samuel Bowles, James Boyce, Sylvia Brandt, Mike Price, Scott Shonkwiler, and seminar participants at the University of Tennessee, Knoxville, the University of Maryland, College Park, the University of Wyoming, and the University of Colorado, Boulder.

# Analyzing Data from Social Dilemma Experiments: A Bayesian Comparison of Parametric Estimators

## Abstract:

Observed choices in Social Dilemma games usually take the form of bounded integers. We propose a doubly-truncated count data framework to process such data. We compare this framework to past approaches based on ordered outcomes and truncated continuous densities via Bayesian estimation and model selection techniques, using data from recent field experiments in rural Colombia. We find that all three frameworks (i) support the presence of treatment-specific unobserved heterogeneity in individual decision-making, and (ii) agree on the ranking of regulatory treatment effects. The count data framework exhibits superior efficiency and produces more informative predictive distributions for outcomes of interest. The truncated continuous framework fails to allocate adequate probability mass to boundary outcomes, which are often of pivotal importance in these games.

Keywords: Social Dilemma Games, Hierarchical Modeling, Bayesian Simulation, Common Property Resource, Institutional Interventions

JEL Codes: C11, C24, C52, C93, Q28

## I) Introduction

Economists often resort to experimental settings to study individuals' decision heuristics in the presence of social externalities. There are two popular variants of such "Social Dilemma" (SD) games: Public Good (PG) games and Common Pool Resource (CPR) games. In the first case subjects need to decide how much of an initial endowment to allocate to a public good (e.g. Andreoni, 1993,1995, Ledyard, 1995, Fehr and Gächter, 2000), while in the second case players need to choose how much to extract from a shared resource (e.g. Ostmann, 1998, Cardenas et al., 2002, Casari and Plott, 2003, Cardenas et al., 2004, Velez et al., forthcoming (a, b)).

The payoff structure for these games is generally designed to create a gap between the game-theoretic equilibrium strategy and the socially optimal course of action. In the bulk of existing applications the researcher's interest centers on the identification of different behavioral motives amongst SD game participants, such as warm-glow, altruism, conditional cooperation, and compliance (e.g. Andreoni, 1995, Anderson et al., 1998, Brandts and Schram, 2001, Fischbacher et al., 2001, Kurzban et al., 2001, Velez et al., forthcoming (b)). A second primary focus in SD games is to study the effect of public institutions and policies on players' allocation or extraction decisions (e.g. Cardenas et al., 2000, Vyrastekova and Soest, 2003, Velez et al., forthcoming (a)). Such institutions are mimicked via exogenous treatments added to the game structure, such as quotas, penalties, open communication, and voting mechanisms. Social Dilemma games with focus on policy effects are often times implemented as field experiments. In such experiments the subject pool is by definition closely linked to important real-world aspects of the game, such as the nature of the commodity under consideration, and /or the nature of the choice tasks or trading rules embedded in the experimental framework (Harrison and List, 2004, List, 2006).

In the vast majority of SD-type games participants have to make discrete choices ("bids") about the per-period level of contribution (in the PG case) or extraction (in the CPR case). In the simplest of PG games this choice set reduces to a basic yes/no decision to contribute (e.g. Offerman et al., 1998) . In many other PG games and most CPR games the set of permissible per-period contribution or extraction

levels ranges from zero or one to some upper bound, say  $E_{\max}$ . The value of  $E_{\max}$  is generally in the 5-30 range (e.g. Andreoni, 1993, Palfrey and Prisbrey, 1996, Fehr and Gächter, 2000, Henrich et al., 2005). In many CPR games  $E_{\max}$  is relatively small (say 5-10) to allow for all payoffs associated with a chosen extraction level, conditional on any possible level of combined extraction by others, to be captured in a single-page spreadsheet (e.g. Cardenas et al., 2000; Cardenas et al., 2002; Casari and Plott, 2003; Bardsley and Moffatt, 2007, Velez et al., forthcoming (a, b)).

Statistically speaking this implies that the outcome variable in most SD games is a doubly-bounded integer. This imposes certain limitations on the econometric framework that can be chosen to process data flowing from SD experiments. Past efforts to analyze SD data with multiple bid levels have primarily relied on models based on an underlying continuous latent variable. For example, Palfrey and Prisbrey (1996) process data from a linear PG experiment using an Ordered Probit framework. Other have applied censored Gaussian regression models (e.g. Carpenter, 2004, Bardsley and Moffatt, 2007, Velez et al., forthcoming (a, b)). We are aware of only a single contribution that employs count data modeling to estimate SD data (Ferraro and Vossler, 2007).

None of these existing econometric approaches fully capture *all* statistical limitations of the dependent variable. Ordered outcomes models essentially ignore the cardinal nature of increasing bid amounts. As a result they require the estimation of an entire set of threshold parameters along with the coefficients needed to reconstruct the latent outcome function. Continuous models, such as the two-limit Tobit (e.g. Carpenter, 2004, Bardsley and Moffatt, 2007) allocate probability mass to outcomes that are not included in the action space of the game. Furthermore, the standard arguments in support of censoring do not apply in this case, since boundary observations in SD experiments do not arise due to non-observability of outcomes outside the bid range. Ferraro and Vossler's (2007) count data approach captures the cardinal integer nature of the dependent variable, but ignores its boundary limitations.

Another important consideration in the analysis of data from SD games is the likely presence of unobserved individual heterogeneity in underlying beliefs and preferences, which can, at least in part,

drive observed differences in decisions and outcomes. The group structure of the experimental setup leaves considerable room for forces such as altruism, inequity aversion, and reputation-building to take root and affect observed outcomes. There is considerable evidence in existing research that these effects exist and that players subscribe to them in heterogeneous fashion (e.g. Palfrey and Prisbrey, 1996, 1997, Anderson et al., 1998, Casari and Plott, 2003). In SD games with external policy components additional heterogeneity effects can arise from differences in how subjects react to prescribed treatments. While this type of “institutional heterogeneity” has been examined to a lesser degree in the existing literature we find strong evidence for this effect in our empirical application.

To date only a limited number of existing contributions have explicitly addressed unobserved individual heterogeneity in their econometric estimation of SD data. Palfrey and Prisbrey (1996) employ a nonparametric individual-level analysis to examine players’ deviation from outcomes prescribed by rational behavior in a voluntary contributions game. Palfrey and Prisbrey (1997) take a closer look at individual-level warm glow effects by incorporating individual fixed effects into a Probit model of voluntary contribution decisions. Messer et al. (2005) use a random-effects regression model to capture unobserved heterogeneity in their analysis of producers’ voluntary contributions to generic advertising. Bardsley and Moffatt (2007) present a finite mixture model that estimates behavior-class probabilities for each individual based on observed choices and individual characteristics in a PG experiment.

In this study we propose a fully parametric econometric framework for the analysis of SD data that captures both the bounded integer nature of the dependent variable and unobserved subject heterogeneity.<sup>1</sup> We label this specification the Hierarchical Doubly-Truncated Poisson (HDTP). We compare this framework to the types that have been used in the past to process such data. Specifically, we contrast results flowing from the HDTP to those produced by a Hierarchical Doubly-Truncated Normal (HDTN) model and a Hierarchical Ordered Probit (HOP) specification. We also estimate non-hierarchical versions of each framework to examine the comparative importance of controlling for unobserved heterogeneity.

We use a Bayesian estimation approach for all specifications. This allows for maximum flexibility in model comparison within and across estimation frameworks via marginal likelihoods and Bayes Factors. We apply our modeling frameworks to recently collected data from a CPR field experiment in three artisanal fishing communities in Colombia. Several key findings flow from this analysis: (i) the marginal likelihood of *any* HDTP sub-model far exceeds the marginal likelihood of the most likely sub-model produced by the other two frameworks, (ii) within each framework, sub-models with hierarchical treatment effects are far more likely than models without treatment indicators and models with fixed treatment effects, (iii) the most likely sub-model within a given framework varies over frameworks, (iv) the HDTP is far more efficient than the other two frameworks as judged by posterior standard deviations for sub-model parameters, and (v) the HDTP produces a more refined and realistic predictive posterior distribution of expected extraction levels than the other two approaches.

Nonetheless, our results also indicate that all three frameworks concur in their ranking of marginal treatment effects and the heterogeneous noise surrounding these effects. In summary, for our SD application we find that the importance of allowing for heterogeneity in treatment effects extends across all frameworks, while the importance of framework choice largely depends on specific research objectives. From an econometric perspective this is, to our knowledge, the first application of a HOP in a Bayesian setting, and the first application of a HDTP model in the broader empirical economics literature.

The remainder of this manuscript is structured as follows: The next section highlights the econometric properties of each estimation framework, Section III describes the field experiment, illustrates the empirical implementation of the three frameworks, and discusses estimation results. Concluding remarks are given in Section IV.

### III) Econometric Frameworks

For consistency with our empirical application we will cast our discussion in the context of a CPR game where each player  $i = 1 \dots n$  in each of  $p = 1 \dots P$  repetitions of the game (“periods”) has to choose a

level of resource extraction or “harvest”  $y_{ip}$  from a given integer range, i.e.

$y_{ip} \in \{E_{\min}, E_{\min} + 1, \dots, E_{\max} - 1, E_{\max}\}$ .<sup>2</sup> In each period players may be exposed to one of several

possible exogenous policy treatments,  $t = 1 \dots T$ , such as harvest quotas with different levels of enforcement, the ability to punish deviations from socially efficient outcomes, or simply the opportunity to communicate prior to choosing harvest levels.

Each modeling framework builds on a conditional mean function of the form

$$y_{ip} = h(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i + \varepsilon_{ip}) \quad (1)$$

where  $y_{ip}$  is the harvest level chosen by individual  $i$  in period  $p$ ,  $\mathbf{x}_i$  is a vector of period-invariant individual characteristics,  $\mathbf{h}_{ip}$  is a vector of treatment indicators that may change over periods and subjects,  $\boldsymbol{\beta}$  denotes a vector of fixed coefficients,  $\boldsymbol{\gamma}_i$  comprises individual-specific random coefficients, and  $\varepsilon_{ip}$  is an i.i.d. error term with zero mean. The random coefficients capture unobserved subject heterogeneity in the effect of experimental treatments on harvest. In all frameworks we model them to follow a multivariate normal distribution with mean  $\boldsymbol{\gamma}$  and variance matrix  $\boldsymbol{\Sigma}$ , i.e.  $\boldsymbol{\gamma}_i \sim mvn(\boldsymbol{\gamma}, \boldsymbol{\Sigma})$ . The additive idiosyncratic error captures additional randomness in observed behavior, for example due to errors in maneuvering through the payoff table.

#### *The Hierarchical Doubly-Truncated Normal (HDTN) Framework*

The HDTN framework corresponds closely to Bardsley and Moffatt's (2007) 2-Limit Tobit model and the Random Effects Tobit specifications used in Carpenter (2004), and in Velez et al., forthcoming (a, b). For this framework the conditional mean function in (1) takes a simple linear form, and the idiosyncratic error  $\varepsilon_{ip}$  follows a normal density with variance  $\sigma^2$ . Observed harvest levels ( $y_{ip}$ ) are mapped to an underlying latent construct of “desired harvest” ( $y_{ip}^*$ ) via the following relationships:

$$\begin{aligned}
y_{ip} &= E_{\min} & \text{if} & & y_{ip}^* \leq E_{\min} \\
y_{ip} &= y_{ip}^* & \text{if} & & E_{\min} < y_{ip}^* < E_{\max} \\
y_{ip} &= E_{\max} & \text{if} & & y_{ip}^* \geq E_{\max}
\end{aligned} \tag{2}$$

The likelihood function for the full sample is given by

$$p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \sigma^2) = \prod_{i=1}^n \int_{\boldsymbol{\gamma}_i} \prod_{p=1}^P \left( \begin{array}{l} \Phi\left(\frac{E_{\min} - (\omega_{ip})}{\sigma}\right) I(y_{ip} = E_{\min}) + \\ 1 - \Phi\left(\frac{E_{\max} - (\omega_{ip})}{\sigma}\right) I(y_{ip} = E_{\max}) + \\ f(y_{ip} | \omega_{ip}, \sigma^2) I(E_{\min} < y_{ip} < E_{\max}) \end{array} \right) f(\boldsymbol{\gamma}_i | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) d\boldsymbol{\gamma}_i \tag{3}$$

where  $\omega_{ip} = \mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i$ ,  $I(\cdot)$  is an indicator function,  $\Phi$  denotes the standard normal cumulative distribution function, and  $f(\cdot | a, b)$  denotes the (uni - or multivariate) normal density with first and second moments given by  $a$  and  $b$ , respectively. The dimension of the integral in (3) corresponds to the length of vector  $\boldsymbol{\gamma}_i$ .

The Bayesian estimation of censored regression models is discussed inter alia in Koop (2004, Ch. 9) and Koop et al. (2007, Ch. 14). We follow these expositions and specify multi-normal priors for the fixed effects and an inverse gamma prior for  $\sigma^2$  i.e.  $\boldsymbol{\beta} \sim mvn(\boldsymbol{\mu}_\beta, \mathbf{V}_\beta)$  and  $\sigma_\varepsilon^2 \sim ig(\eta_0, \kappa_0)$ , where  $mvn(\cdot)$  indicates the multivariate normal density and  $ig(a, b)$  denotes the inverse gamma distribution with shape parameter  $a$  and scale  $b$ . We use the same densities for the hyper-priors of the hierarchical mean vector and the diagonal elements of the hierarchical variance matrix, respectively, i.e.

$$\boldsymbol{\gamma} \sim mvn(\boldsymbol{\mu}_\gamma, \mathbf{V}_\gamma), \boldsymbol{\Sigma}_{jj} \sim ig(\nu_0, \varphi_0), j = 1 \dots k_r, \text{ where } k_r \text{ is the number of treatment indicators in } \boldsymbol{\gamma}_i.^3$$

As for other limited dependent variable models the estimation of the HDTN is facilitated by treating latent harvest  $y_{ip}^*$  as augmented data (Tanner and Wong, 1987) that are drawn along with the

actual model parameters in the posterior simulator. Using a Gibbs Sampler (GS) we draw consecutively and repeatedly from the following set of conditional posterior densities:

$$\begin{aligned}
 & p(\boldsymbol{\beta} | \mathbf{y}^*, \mathbf{X}, \mathbf{H}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\Sigma}), \quad p(\boldsymbol{\gamma} | \mathbf{y}^*, \mathbf{X}, \mathbf{H}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\Sigma}), \quad p(\gamma_i | y_i^*, \mathbf{X}_i, \mathbf{H}_i, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\Sigma}, \boldsymbol{\gamma}), i=1 \cdots n, \\
 & p(\boldsymbol{\Sigma}_{jj} | \boldsymbol{\gamma}, \gamma_i), i=1 \cdots n, j=1 \cdots k_r, \quad p(\sigma^2 | \mathbf{y}^*, \mathbf{X}, \mathbf{H}, \boldsymbol{\beta}, \boldsymbol{\gamma}_i), i=1 \cdots n \quad \text{and} \\
 & p(\mathbf{y}^* | \mathbf{y}, \mathbf{X}, \mathbf{T}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\gamma}_i), i=1 \cdots n,
 \end{aligned} \tag{4}$$

where matrices  $\mathbf{X}$  and  $\mathbf{H}$  comprise, respectively, vectors  $\mathbf{x}_i$  and  $\mathbf{h}_{ip}$ , for all individuals and periods, and matrices  $\mathbf{X}_i$  and  $\mathbf{H}_i$  perform the same function for a given individual across all periods. As indicated in (4) both the set of fixed coefficients and the mean vector for the random coefficients are drawn *unconditional* on  $\gamma_i$  to improve the efficiency of the sampler.<sup>4</sup> The detailed expressions for all conditional posterior densities are given in Appendix A.

As discussed e.g. in Gelman et al. (2004) after a sufficient number of repetitions the conditional draws of  $\boldsymbol{\beta}$ ,  $\boldsymbol{\gamma}$ ,  $\boldsymbol{\Sigma}$ , and  $\sigma^2$  will converge to the joint posterior distribution  $p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \sigma^2 | \mathbf{y}, \mathbf{X}, \mathbf{H})$ . Furthermore, each series of draws for a given subset of parameters by itself represents draws from the respective marginal posterior. Econometric inference can then be conducted through an inspection of the shape and / or the moments of these marginal distributions.

#### *The Hierarchical Ordered Probit (HOP) Framework*

The HOP framework extends Palfrey and Prisbrey's (1996) Ordered Probit approach to analyze SD data by incorporating individual heterogeneity via a hierarchical layer. As in the HDTN framework we interpret observed chosen harvest levels as manifestations of an underlying latent “desired harvest” function. The latent conditional mean function and stochastic model elements are as for the HDTN with the added restriction of  $\sigma^2 = 1$  for identification purposes.

Observed effort is mapped to the latent effort function via the following probability brackets:

$$\begin{aligned}
pr(y_{ip} = E_{\min}) &= pr(-\infty < y_{ip}^* \leq 0) = pr(-\infty < \varepsilon_{igp} \leq -\omega_{ip}) = \Phi(-\omega_{ip}) \\
pr(y_{ip} = E_{\min} + 1) &= pr(0 < y_{ip}^* \leq c_1) = \Phi(c_1 - \omega_{ip}) - \Phi(-\omega_{ip}) \\
pr(y_{ip} = E_{\min} + 2) &= pr(c_1 < y_{ip}^* \leq c_2) = \Phi(c_2 - \omega_{ip}) - \Phi(c_1 - \omega_{ip}) \\
&\vdots \\
pr(y_{ip} = E_{\max} - 1) &= pr(c_{E_{\max}-3} < y_{ip}^* \leq c_{E_{\max}-2}) = \Phi(c_{E_{\max}-2} - \omega_{ip}) - \Phi(c_{E_{\max}-3} - \omega_{ip}) \\
pr(y_{ip} = E_{\max}) &= pr(c_{E_{\max}-2} < y_{ip}^* \leq \infty) = 1 - \Phi(c_{E_{\max}-2} - \omega_{ip}),
\end{aligned} \tag{5}$$

where as before  $\omega_{ip} = \mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i$ , and the symbol  $\Phi$  denotes the standard normal cumulative distribution function. The  $c$ -terms in (5) are bin-thresholds that are estimated along with the other model parameters. Following standard convention we impose the additional identification restriction of setting the first threshold equal to zero, as evident from the first two lines in (5). We will index the remaining thresholds as  $c_b, b = 1 \dots B$ , where  $B = E_{\max} - 2$ .

The likelihood function for the full model can be generically written as

$$p(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \mathbf{c}) = \prod_{i=1}^n \int_{\boldsymbol{\gamma}_i} \left( \prod_{p=1}^P \left( \sum_{e=E_{\min}}^{E_{\max}} pr(y_{ip} = e) I(y_{ip} = e) \right) f(\boldsymbol{\gamma}_i | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) \right) d\boldsymbol{\gamma}_i \tag{6}$$

where  $\mathbf{c}$  comprises all threshold terms and  $pr(y_{ip} = e)$  is given in (5) for any permissible extraction level  $e$ .

Bayesian Ordered Probit models or model components have recently been employed by Huang and Lin (2006), Li and Tobias (2006), and Li and Tobias (2008). To assure a rapid mixing and to reduce serial correlation of parameter draws in our Bayesian simulation routine we follow Nandram and Chen (1996) and Li and Tobias (2008) and re-parameterize the HOP model using the inverse of the highest bin threshold as follows:

$$\begin{aligned}
\tilde{y}_{ip}^* &= \mathbf{x}'_i \tilde{\boldsymbol{\beta}} + \mathbf{h}'_{ip} \tilde{\boldsymbol{\gamma}}_i + \tilde{\varepsilon}_{ip} & \tilde{\varepsilon}_{ip} &\sim n(0, \delta^2), \quad \text{where} \\
\tilde{\boldsymbol{\beta}} &= \delta \boldsymbol{\beta}, \quad \tilde{\boldsymbol{\gamma}}_i = \delta \boldsymbol{\gamma}_i, \quad \tilde{\varepsilon}_{ip} = \delta \varepsilon_{ip}, \quad \text{and} & \delta &= \frac{1}{c_B}
\end{aligned} \tag{7}$$

This implies the following re-parameterized values for the hierarchical moments and bin-thresholds:

$$\tilde{\boldsymbol{\gamma}}_i \sim mvn(\tilde{\boldsymbol{\gamma}}, \tilde{\boldsymbol{\Sigma}}), \quad \tilde{\boldsymbol{\gamma}} = \delta \boldsymbol{\gamma}, \quad \tilde{\boldsymbol{\Sigma}} = \delta^2 \boldsymbol{\Sigma}, \quad \tilde{c}_b = \delta c_b, b=1 \dots B, \quad \text{and} \quad \tilde{c}_B = 1. \quad (8)$$

The Bayesian estimation routine is specified in terms of these transformed parameters. However, the actual posterior parameter values *collected* during simulation are the original parameters of interest. In analogy to the HDTN we specify the following priors and hyper-priors for the transformed parameters:

$$\tilde{\boldsymbol{\beta}}_i \sim mvn(\boldsymbol{\mu}_\beta, \mathbf{V}_\beta), \quad \delta^2 \sim ig(\eta_0, \kappa_0), \quad \tilde{\boldsymbol{\gamma}} \sim mvn(\boldsymbol{\mu}_\gamma, \mathbf{V}_\gamma), \quad \tilde{\boldsymbol{\Sigma}}_{jj} \sim ig(\nu_0, \varrho_0), \quad j=1 \dots k_r. \quad (9)$$

It is common to assign diffuse (or “improper”) priors for the threshold parameters in a Bayesian Ordered Probit framework (e.g. Li and Tobias, 2006, Li and Tobias, 2008, Koop, 2004, Ch. 9). While this strategy is computationally convenient it preempts a proper comparison of models within the HOP framework with models in other frameworks via marginal likelihoods and Bayes Factors (e.g. Kass and Raftery, 1995). We thus prefer a proper prior for these elements and specify

$$\tilde{c}_b \sim n(\mu_c, V_c) I(\tilde{c}_b \in [0, 1]), \quad (10)$$

i.e. we a priori let threshold parameters follow a normal density with mean  $\mu_c$  and variance  $V_c$ , truncated to the support defined by the lowest and highest re-parameterized threshold values.

As for the HDTN the Bayesian posterior simulator for the HOP specification is best implemented by augmenting the set of parameters of interest with actual draws of the latent data vector  $\tilde{\mathbf{y}}^*$  (Tanner and Wong, 1987, Albert and Chib, 1993). We use a Gibbs Sampler (GS) to draw consecutively and repeatedly from the following set of conditional posterior densities:

$$\begin{aligned} & p(\tilde{\boldsymbol{\beta}} | \tilde{\mathbf{y}}^*, \mathbf{X}, \mathbf{H}, \tilde{\boldsymbol{\gamma}}, \delta^2, \tilde{\boldsymbol{\Sigma}}), \quad p(\tilde{\boldsymbol{\gamma}} | \tilde{\mathbf{y}}^*, \mathbf{X}, \mathbf{H}, \tilde{\boldsymbol{\beta}}, \delta^2, \tilde{\boldsymbol{\Sigma}}), \quad p(\delta^2 | \tilde{\mathbf{y}}^*, \mathbf{X}, \mathbf{H}, \tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\gamma}}_i), i=1 \dots n \\ & p(\tilde{\mathbf{c}} | \mathbf{y}, \tilde{\mathbf{y}}^*, \mathbf{X}, \mathbf{H}, \tilde{\boldsymbol{\beta}}, \delta^2, \tilde{\boldsymbol{\gamma}}_i), i=1 \dots n, \quad p(\tilde{\boldsymbol{\Sigma}}_{jj} | \tilde{\boldsymbol{\gamma}}, \tilde{\boldsymbol{\gamma}}_i), i=1 \dots n, j=1 \dots k_r, \\ & p(\tilde{\mathbf{y}}^* | \mathbf{y}, \mathbf{X}, \mathbf{T}, \tilde{\boldsymbol{\beta}}, \delta^2, \tilde{\boldsymbol{\gamma}}_i, \tilde{\mathbf{c}}), i=1 \dots n \quad \text{and} \quad p(\tilde{\boldsymbol{\gamma}}_i | \tilde{\mathbf{y}}_i^*, \mathbf{X}_i, \mathbf{H}_i, \tilde{\boldsymbol{\beta}}, \delta^2, \tilde{\boldsymbol{\Sigma}}, \tilde{\boldsymbol{\gamma}}), i=1 \dots n. \end{aligned} \quad (11)$$

where  $\tilde{\mathbf{c}}$  comprises all re-parameterized threshold parameters. As for the HDTN the set of fixed coefficients and the mean vector for the random coefficients are drawn unconditional on individual vectors  $\boldsymbol{\gamma}_i$ . The draws of transformed thresholds in  $\tilde{\mathbf{c}}$  require a Metropolis-Hastings routine within the

GS, along the lines discussed in Li and Tobias (2008). The detailed expressions for all conditional posterior densities are given in Appendix B. After a sufficient number of repetitions the conditional draws of  $\boldsymbol{\beta}$ ,  $\boldsymbol{\gamma}$ ,  $\boldsymbol{\Sigma}$ , and  $\mathbf{c}$  (restored to their original form) will converge to the joint posterior distribution

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \mathbf{c} | \mathbf{y}, \mathbf{X}, \mathbf{H}).$$

*The Hierarchical Doubly-Truncated Poisson Model (HDTP)*

Based on considerations of parsimony and flexibility we choose a Poisson kernel to form the foundation of our count data specification to model the integer extraction levels. A general discussion of Poisson models is given e.g. in Cameron and Trivedi (1998) and Winkelmann (2003). The estimation of a hierarchical Poisson model in a Bayesian framework has been illustrated in Chib and Carlin (1999). We extend their specification by adding a double-sided truncation to the Poisson kernel. The resulting density for *observed* effort level  $y_{ip}$  follows as

$$f(y_{ip} | \lambda_{ip}, E_{\min} \leq y_{ip} \leq E_{\max}) = \frac{\exp(-\lambda_{ip}) \lambda_{ip}^{y_{ip}}}{y_{ip}! \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\exp(-\lambda_{ip}) \lambda_{ip}^k}{k!} \right)} = \frac{\lambda_{ip}^{y_{ip}}}{y_{ip}! \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\lambda_{ip}^k}{k!} \right)} \quad \text{where} \quad (12)$$

$$\lambda_{ip} = \exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i) \quad \text{and} \quad \boldsymbol{\gamma}_i \sim mvn(\boldsymbol{\gamma}, \boldsymbol{\Sigma}).$$

In this case the conditional mean or “link” function  $\lambda_{ip}$  denotes the expectation and variance of the *untruncated* density. As opposed to the first two frameworks this function is nonlinear and does not contain an additive error term. The mean and variance of the truncated density, conditional on regressors,  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}_i$ , can be expressed as (see e.g. Cameron and Trivedi, 1998)

$$\begin{aligned}
E(y_{ip} | \lambda_{ip}, E_{\min} \leq y_{ip} \leq E_{\max}) &= \lambda_{ip} \left( 1 + \frac{f(E_{\min} - 1) - f(E_{\max})}{F(E_{\min}, E_{\max})} \right) \\
V(y_{ip} | \lambda_{ip}, E_{\min} \leq y_{ip} \leq E_{\max}) &= \lambda_{ip} \left( 1 - \frac{(\lambda_{ip} + E_{\min}) f(E_{\min} - 1)}{F(E_{\min}, E_{\max})} - \frac{(E_{\max} + 1 - \lambda_{ip}) f(E_{\max})}{F(E_{\min}, E_{\max})} - \lambda_{ip} \frac{(f(E_{\min} - 1) - f(E_{\max}))^2}{F(E_{\min}, E_{\max})^2} \right)
\end{aligned} \tag{13}$$

where  $f(k) = \frac{\exp(-\lambda_{ip}) \lambda_{ip}^k}{k!}$  and  $F(E_{\min}, E_{\max}) = \sum_{k=E_{\min}}^{E_{\max}} f(k)$ ,  $E_{\min} > 0$

Thus, the magnitude of the truncated expectation in relation to the untruncated mean depends on the comparative magnitude of the untruncated densities at  $E_{\min} - 1$  and  $E_{\max}$ . The truncated variance will be unambiguously smaller than the untruncated version as long as  $E_{\max} + 1 \geq \lambda_{ip}$ . Therefore, for the dual reasons of truncation and parameter heterogeneity the mean-variance equality constraint of the standard Poisson density no longer holds. Furthermore, as for the HDTP and the HOP, the random coefficients in  $\lambda_{ip}$  introduce correlation between decisions made by a given individual over the  $P$  periods of the game.

The likelihood for the HDTP is given as

$$p(y_{ip} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, E_{\min} \leq y_{ip} \leq E_{\max}) = \prod_{i=1}^n \int_{\boldsymbol{\gamma}_i} \left( \prod_{p=1}^P \frac{\lambda_{ip}^{y_{ip}}}{y_{ip}! \left( \sum_{k=L}^U \frac{\lambda_{ip}^k}{k!} \right)} \right) f(\boldsymbol{\gamma}_i | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) d\boldsymbol{\gamma}_i \tag{14}$$

As for the HDTP and HOP we specify multivariate normal priors for  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}$  and an inverse gamma prior for the elements of  $\boldsymbol{\Sigma}$ , i.e.

$$\boldsymbol{\beta} \sim mvn(\boldsymbol{\mu}_{\boldsymbol{\beta}}, \mathbf{V}_{\boldsymbol{\beta}}), \quad \boldsymbol{\gamma} \sim mvn(\boldsymbol{\mu}_{\boldsymbol{\gamma}}, \mathbf{V}_{\boldsymbol{\gamma}}), \quad \Sigma_{jj} \sim ig(v_0, \varphi_0), \quad j = 1 \dots k_r \tag{15}$$

The posterior simulator draws from the following conditional probabilities:

$$\begin{aligned}
p(\boldsymbol{\beta} | \mathbf{y}, \mathbf{X}, \mathbf{H}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}_i), i = 1 \dots n, \quad p(\boldsymbol{\gamma} | \mathbf{y}, \mathbf{X}, \mathbf{H}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}_i), i = 1 \dots n, \\
p(\boldsymbol{\Sigma}_{jj} | \boldsymbol{\gamma}, \boldsymbol{\gamma}_i), i = 1 \dots n, j = 1 \dots k_r, \quad \text{and} \quad p(\boldsymbol{\gamma}_i | \mathbf{y}_i, \mathbf{X}_i, \mathbf{H}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}), i = 1 \dots n
\end{aligned} \tag{16}$$

Posterior inference is based on the marginals of the joint posterior distribution  $p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma} | \mathbf{y}, \mathbf{X}, \mathbf{H})$ . In comparison to the HDTP and HOP the posterior simulator for the HDTP does not require draws of threshold parameters or latent dependent variables. However, its computational implementation is more complex as it requires Metropolis-Hastings (MH) steps for draws of  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}_i$ . Chib and Carlin (1999) offer several possible versions for these MH subroutines. We achieved the most efficient results by choosing a reflected tailored proposal density for draws of  $\boldsymbol{\beta}$ , and a tailored proposal density with an Accept-Reject (A/R) primer for draws of  $\boldsymbol{\gamma}_i$ . The detailed steps of this posterior simulator are given in Appendix C.

### *Marginal Likelihood and Model Selection*

As described in more detail in the next section we estimate an identical set of  $S$  different *sub-models* for each estimation framework. These sub-models are distinguished by the contents and hierarchical properties of treatment vector  $\mathbf{h}_{ip}$  in (1). Let a *model* be defined as any combination of framework and sub-model, and let  $M$  denote the total number of estimated models. Thus, given our three estimation frameworks we have  $M = 3S$ . Each model is associated with a posterior model probability, given as

$$p(m | \mathbf{y}) = \frac{p(\mathbf{y} | m) p(m)}{p(\mathbf{y})} \quad (17)$$

where  $p(\mathbf{y} | m)$  is the model- conditioned marginal likelihood,  $p(m)$  indicates the prior model probability, and  $p(\mathbf{y})$  denotes the unconditional marginal likelihood, i.e. the probability that  $\mathbf{y}$  was generated by *any* of the considered models. For ease of exposition we omit any explicit conditioning on explanatory variables.

Loosely speaking the posterior model probability indicates how likely the observed data (i.e. the vector of dependent observations,  $\mathbf{y}$ ) were generated by model  $m$ . As discussed e.g. in Koop (2004, Ch.1) and evident from (17) under equal model priors the ratio of model probabilities for two competing

specifications reduces to the ratio of marginal likelihoods. This ratio is commonly referred to as Bayes Factors (*BF*) (e.g. Kass and Raftery, 1995). It captures the relative probability that either model is the correct specification given the underlying data. As discussed e.g. in Geweke (2005) it is not necessary for two models to be nested for the *BF* to be applicable. This makes the *BF* an extremely versatile tool for model comparison and selection, as we illustrate in our empirical section.

To compute *BF*s one must first evaluate the marginal likelihood for each of the two models under consideration. Dropping model-conditioning for simplicity, the marginal likelihood for any given model can be expressed as

$$p(\mathbf{y}) = \frac{p(\mathbf{y} | \boldsymbol{\theta}_m) p(\boldsymbol{\theta}_m)}{p(\boldsymbol{\theta}_m | \mathbf{y})} \quad (18)$$

where  $\boldsymbol{\theta}_m$  is the vector of parameters corresponding to model  $m$ ,  $p(\mathbf{y} | \boldsymbol{\theta}_m)$  is the sample likelihood,  $p(\boldsymbol{\theta}_m)$  denotes the set of parameter priors, and  $p(\boldsymbol{\theta}_m | \mathbf{y})$  is the posterior distribution of model parameters. The marginal likelihood is often also referred to as “prior predictive distribution” (e.g. Lancaster, 2004, Ch. 2). It describes what new data are expected to look like before they are collected, given the researcher’s choice of likelihood function and prior. Since the left hand side of (18) does not contain  $\boldsymbol{\theta}_m$ , the relationship must hold for any choice of  $\boldsymbol{\theta}_m$ . This property is deemed “marginal likelihood identity” in Chib (1995). In practice,  $p(\mathbf{y})$  is usually evaluated at a point of high posterior density for  $\boldsymbol{\theta}_m$ . Since an analytical expression for  $p(\mathbf{y})$  does not exist for any of our frameworks and sub-models we simulate this value following the methods described in Chib (1995) for the HDTN and the technique outlined in Chib and Jeliazkov (2001) for the HOP and the HDTP.<sup>5</sup>

### *Posterior Predictions*

For each estimation framework and sub-model the posterior sampler generates  $r=1 \dots R$  draws of parameters. To derive posterior predictive distributions (PPDs) of extraction probabilities or extraction levels these draws need to be combined with specific settings for individual characteristics and

experimental treatments. We will denote these as  $\mathbf{x}_f$  and  $\mathbf{h}_f$ , respectively. Generically, the PPD for any function of individual characteristics and experimental treatments,  $g(\mathbf{x}_f, \mathbf{h}_f)$ , conditional only on model  $m$  can be expressed as

$$p(g(\mathbf{x}_f, \mathbf{h}_f) | m) = \int_{\boldsymbol{\theta}} \left\{ \int_{\gamma_i} g(\mathbf{x}_f, \mathbf{h}_f | \gamma_i, \boldsymbol{\theta}) f(\gamma_i | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) d\gamma_i \right\} p(\boldsymbol{\theta} | \mathbf{y}, \mathbf{X}, \mathbf{H}, m) d\boldsymbol{\theta} \quad (19)$$

where the exact form of  $g(\cdot)$  depends on the desired predictive construct and the estimation framework, and  $\boldsymbol{\theta}$  comprises all model parameters. For example,  $g(\cdot)$  can denote expected harvest, a specific extraction level, or the probability that extraction falls between two specific bounds. The detailed expressions for these predictive constructs and the practical implementation steps to derive them are given in Appendix D.

#### IV) Empirical Application

##### *Data*

To illustrate our estimation framework we use data from CPR field experiments conducted in 2004 in three rural communities of Colombia. Each community relies primarily on artisanal fishing. In each area 60 subjects were recruited to play 20 rounds of a CPR game in groups of five players. In each round subjects had to decide to harvest between 1 and 9 units of the resource. While the resource was loosely described as a “fishery” the units of extraction were not explicitly defined. Instead all participants received an identical payoff table in which each cell constitutes the intersection between harvest level chosen by the individual and the combined harvest level of all other group members. The table is shown in Appendix E. Payoffs were derived using the following underlying profit function

$$\pi_i = a + by_i - c(y_i + Y_{-i}) - dy_i(y_i + Y_{-i}) \quad \text{with} \quad (20)$$

$$a = 858.75, \quad b = 130.625, \quad c = 15.125, \quad \text{and} \quad d = 2.75,$$

where  $\pi_i$  is the per-period payoff,  $y_i$  is the harvest level chosen by subject  $i$ , and  $Y_{-i}$  is the combined harvest by all other members in the group. The profit function thus incorporates both a stock externality

and a cost externality, as is typical for CPR problems for fishing applications. The socially optimal per – period harvest and the pure strategy Nash equilibrium lie at two and seven units, respectively.

The game was administered under different institutional “treatments”. For the first 10 rounds all of the 12 groups in a given community played an identical “Open Access” version of the game without government intervention or communication (although with an upper bound of nine units of harvest). We will label this version “Treatment 1” ( $T_1$ ). In the second set of 10 rounds one of the following three treatments was implemented for sets of four groups each per community: (i) “Treatment 2” ( $T_2$ ), a harvest quota of 2 units, with a 10% audit probability and a low penalty (27 pesos for each unit exceeding 2)<sup>6</sup>, (ii) “Treatment 3” ( $T_3$ ), a harvest quota of 2 units, with a 10% audit probability and a medium penalty (165 pesos for each unit exceeding 2), and (iii) “Treatment 4” ( $T_4$ ), the opportunity to openly communicate with the other group members prior to each round of decision-making.<sup>7</sup> Further details on the implementation of the experiment are given in Velez et al., forthcoming (a).

As described in Velez et al. (forthcoming (a)) there exist pronounced cultural, socio-economic, and institutional differences between the three communities. For example, the first community, Ensenada de Tumaco on the Pacific coast, has a culturally homogeneous population for whom shrimp fishing is the primary economic activity. There is a long-standing tradition of government regulations for this type of fishery. In contrast, residents of La Dorada in Colombia’s interior depend primarily on fish species from the Magdalena River watershed. There is only a weak presence of regulatory mechanisms for this fishery. Instead local communities have traditionally relied on informal associations to manage these resources. The third community, located near the city of Santa Marta on the Caribbean coast, has the most ethnically diverse population of the three regions. Also, fishing is somewhat less important than in the other two areas, and few fishing rules, formal or informal, appear to be recognized or observed in this community. Compared to 94% for Tumaco and 87% for La Dorada, only 69% of participants in the Caribbean CPR game indicated that fishing was their main economic activity. The remaining participants were primarily employed in the agricultural sector. In the following we will label the three areas as “Pacific” (PAC),

“Magdalena” (MAG), and “Caribbean” (CAR), respectively. Basic sample statistics for each community are given in Table 1.

### *Specification of Sub-models*

For each estimation framework we specify five sub-models, each with a different conditional mean function. The salient features of these sub-models are captured in the first four columns of Table 2. For all sub-models the vector of individual characteristics  $\mathbf{x}_i$  includes “gender” (1 = female) and “years of education”<sup>8</sup>. Sub-model one ( $M_1$ ) adds only a generic random effect to this set of individual characteristics. Sub-model two ( $M_2$ ) replaces this random effect with fixed treatment indicators for each of the four regulatory interventions. Sub-model three ( $M_3$ ) changes the latter to random indicators. Thus this model examines if heterogeneity in observed decisions manifests itself differently under different treatments, i.e. how much “noise” in decision-making is associated with each treatment. Sub-models four and five capture time trends, perhaps due to learning effects, via a binary indicator (“*period*”) for periods 6-10 vs. periods 1-5 for each 10-period cycle of the experiment. Specifically, sub-model four ( $M_4$ ) expands  $M_2$  by adding an interaction term with *period* for each treatment to the set of fixed effects. Sub-model five ( $M_5$ ), in turn, includes *period* as an additional random effect in lieu of the fixed interactions. Intuitively,  $M_4$  allows time trends to operate systematically differently across treatments, while  $M_5$  stipulates that decision trends are primarily related to unobserved heterogeneity unrelated to treatment. Naturally, this set of sub-models could easily be expanded to allow for additional specifications. However, the five versions describe above suffice to address our main research questions in a parsimonious fashion.

### *Estimation Results*

We estimate all models using the following vague but proper priors:

$\boldsymbol{\mu}_\beta = 0, \mathbf{V}_\beta = 100, \boldsymbol{\mu}_\gamma = 0, \mathbf{V}_\gamma = 100, \eta_0 = \kappa_0 = \nu_0 = \varphi_0 = 1/2$ . The tuner elements in the MH algorithms for

draws of the threshold parameters  $\tilde{c}$  for the HOP are set to achieve a recommended acceptance rate of 0.40 – 0.45 (e.g. Gelman et al., 2004, Ch. 11). All models are estimated using 3000 burn-in draws and 120000 retained draws in the Gibbs Sampler. The decision on the appropriate amount of burn-ins was guided by Geweke's (1992) convergence diagnostic (CD).

Table 2 captures the essential features of each sub-model as well as marginal likelihood ( $mLH$ ) and Bayes Factors ( $BFs$ ), in log form, for all three communities. Within each estimation framework the reported  $BFs$  refer to the ratio of the  $mLH$  of the most likely sub-model to all other specifications. As can be seen from the first block of the table, within the HDTN framework  $M_3$  achieves the highest (i.e. least negative) logged  $mLH$  for the Pacific and Caribbean samples. For Magdalena, sub-model  $M_5$  is virtually as likely as sub-model  $M_3$ . Using the interpretation thresholds for  $BFs$  suggested in Kass and Raftery, (1995), there is “decisive” ( $\log BF > 11.5$ ) evidence that  $M_3$  ( $M_5$  for Magdalena) is more likely to have generated the observed data than the random constant-only model ( $M_1$ ) and the model with fixed treatment effects ( $M_2$ ), and “strong” ( $\log BF > 6.9$ ) evidence that the most likely sub-model better describes the underlying data than the sub-model with fixed treatment-period interactions ( $M_4$ ).

This overall picture is similar but sharper for the HOP framework (second block of rows in Table 2), where the sub-model with random treatment effects and no period-effect ( $M_3$ ) outperforms all other sub-models in a decisive fashion for all three regions. Thus, if one were to employ the HOP framework to analyze these data one would conclude that it is highly improbable that there exist “period” effects, i.e. that observed harvest decisions change between the first and second half of a given version of the SD game.

In contrast, within the HDTP framework sub-model  $M_5$  with a random period effect in addition to random treatment effects emerges as overwhelmingly superior to all other sub-models for all three regions, as is apparent from the last block of rows in the table. In summary, all three frameworks agree in the sense that they allocate higher probabilities to sub-models that allow for unobserved heterogeneity in treatment effects. This is one of the key findings flowing from our analysis. A second key finding is that the frameworks disagree as to the importance of temporal effects in observed decisions. At a general

level this implies that the choice of estimation framework can lead to different preferred specifications with respect to the conditional mean function, and different inferences regarding the effects of regressors on the outcome variable. The natural next step would be to compare the framework-specific “winners” to each other. In a classical framework this would pose somewhat of a dilemma as it would be difficult to rigorously compare sub-models *across* estimation frameworks. In contrast, our Bayesian approach to model comparison via *mLHs* and *BFs* is suitable for the assessment of relative probabilities for any two models that share the same dependent variable, regardless of the underlying estimation framework.

Table 3 provides such a comparison. It assigns a block of three rows to each of the three regions. Each row, in turn, corresponds to the best model for each of the three frameworks. As mentioned above we use the simulation method suggested by Chib (1995) and Chib and Jeliazkov (2001) to approximate the marginal likelihood. We evaluate the *mLH* in log form at the posterior mean for each model.

Building on (18), we thus have

$$\log(p(\mathbf{y})) = \log(p(\mathbf{y} | \bar{\boldsymbol{\theta}}_m)) + \left( \log(p(\bar{\boldsymbol{\theta}}_m)) - \log(p(\bar{\boldsymbol{\theta}}_m | \mathbf{y})) \right) \quad (21)$$

where  $\bar{\boldsymbol{\theta}}$  denotes the posterior mean of all model parameters. Thus, the log *mLH* for a given model will be high if the sample likelihood is high at the posterior mean (i.e. if the chosen likelihood function is supported by the data), and if the posterior density at  $\bar{\boldsymbol{\theta}}$  is not substantially larger than the prior ordinate at  $\bar{\boldsymbol{\theta}}$ , controlling for the information content in  $\mathbf{y}$ . The latter condition implies that the prior was *ex ante* well-chosen. Under vague priors the difference between posterior and prior ordinate in (21) will likely be considerable, especially if the parameter space is large. In that case a high *mLH* score will largely rest on the appropriateness of the likelihood function.

To highlight these disparate effects Table 3 depicts the prior density and the sample likelihood along with the *mLH* for each model. As can be seen from the fourth column in the table there are only minor differences in prior densities, evaluated at the posterior mean, across models within each region. This is not surprising given the comparable parameter spaces and the identical (non-informative) priors chosen for equivalent parameters. As is clear from the next two columns, the observed large differences

in  $mLH$  are largely driven by sample likelihood values. The main insight from the table is that the HDTP, represented by sub-model  $M_5$ , produces by far the highest log-likelihood values and  $mLH$  scores for all three regions. Expressed in  $BFs$  (last column of Table 3), this model receives overwhelming support over the top sub-models produced by the other two frameworks. It should also be noted that the best sub-model identified under the HOP framework, while clearly inferior to the HDTP sub-model, receives still vastly higher posterior probability than the best model flowing from the HDTN framework.

This pattern essentially reflects the degree of efficiency (or lack thereof) with which the three frameworks process the collected data. It leaves little doubt that the HDTP is more suitable to analyze these data than the other two approaches. However, it remains to be seen if these pronounced differences in framework efficiency and model probabilities translate into different inferential outcomes with respect to key research questions.

To examine this issue we first compare posterior results for all parameters for the most likely sub-model within each framework. For simplicity and ease of comparison we choose  $M_3$  for the HDTN for all three regions, even though  $M_5$  receives slightly higher posterior mass for the Magdalena experiment within that framework (see Table 2). Table 4 shows the posterior mean and standard deviation for all parameters and regions.<sup>9</sup> The first triplet of columns corresponds to sub-model  $M_3$  for the HDTN, the second triplet to  $M_3$  for the HOP, and the last three columns to sub-model  $M_5$  within the HDTP framework. The first block of rows shows results for the fixed effects, i.e. the elements of  $\beta$ , the second block corresponds to the means of random effects, i.e. the elements of  $\gamma$ , and the third block gives results for the variances contained in  $\Sigma$ .

The estimated posterior means for the fixed effects “female” and “education” can be interpreted as the marginal effects of these regressors on latent extraction for the HDTN and the HOP, and on the logged (un-truncated) expected extraction for the HDTP. The posterior means for expected treatment effects, depicted in the second block of rows in the table, can be loosely interpreted as expected latent harvest levels for a male participant with no schooling for the HDTN and the HOP, and as the log of

expected harvest levels for the same individual for the HDTP.<sup>10</sup> The posterior means for random effect variances, given in the third block of rows in the table, reflect the degree of unobserved heterogeneity surrounding each treatment effect. The posterior standard deviations, provided in parentheses beneath each posterior mean for all parameters, indicate the level of precision with which these effects are estimated.

There are four important results that flow from Table 4: (i) the HDTP produces the smallest posterior standard deviations for all (shared) parameters, followed by the HOP and the HDTN. This reflects the above mentioned ranking in efficiency of data exploitation across the three frameworks. (ii) Within each region, all three frameworks generate the *same relative ranking of expected treatment effects*;<sup>11</sup> (iii) the communication treatment  $T_4$  exhibits considerably smaller heterogeneity noise than the two regulatory interventions based on quotas ( $T_2$  and  $T_3$ ), and (iv) game-period effects on harvest levels, as captured by the HDTP, play a minor role compared to treatment effects and vary across regions.

With respect to finding (ii) we note that, not surprisingly, the posterior means of expected harvest levels are highest for the open access scenario ( $T_1$ ) for all regions and models. For PAC the institutional treatments  $T_2$  and  $T_3$  have a decidedly stronger effect on harvest reduction than communication, with the low-penalty version  $T_2$  producing the lowest extraction levels. This picture is reversed for CAR and MAG, where  $T_4$  induces the lowest mean extraction levels followed by the medium and low penalty regulatory scenario, respectively. Again, this result is intuitively sound when considering that participants in PAC are used to governmental regulation of their fishery and have historically relied less on cooperative conservation efforts.

In contrast, for the participants in MAG who have a history of cooperative governance of their fishery, communication is generally more effective than regulatory institutions. For CAR, in turn, where neither local nor national regulatory institutions are well established (or acknowledged), the strong effect of communication to reduce harvest may largely be attributable to the “trust-building” effect of open discussion, especially given the more transient nature of that society and the relatively large share of non-fishermen in the experiment.<sup>12</sup>

Following up on finding (iii) regarding random effect variances, we first observe that individual heterogeneity vis-à-vis treatment effects appears to be least pronounced for the open access treatment  $T_1$ . However, econometric considerations warrant some caution in interpreting these results. Specifically, the design of the experiment produces three times as many observations ( $n = 600$ ) for decisions under  $T_1$  than for the other three treatments ( $n = 200$  each). This implies that three times more data points are available to update the vague variance prior for  $T_1$  than for the other treatments. Given our overall moderate number of observations it is thus possible that the observed “diminished heterogeneity” with respect to  $T_1$  is at least in part driven by these sample size effects. The remaining variance terms, however, are on equal footing in this regard. Our results clearly indicate that allowing for communication reduces heterogeneous noise in decision making compared to external regulations, i.e. fosters the coordination of individual strategies. To our knowledge this explicit insight into the second order effects of communication is novel in the experimental literature.

With respect to result (iv) noted above, we find that the posterior mean for the expected second-period-effect (last row of the second block for the HDTP) is located slightly below zero for PAC and slightly above zero for CAR and MAG. In addition, the posterior means of random effect variances (last row of the third block for the HDTP) are relatively close to zero. We therefore conclude that changes in harvest strategies that are purely a function of the duration of a given game are at best subtle, and clearly less important than changes in harvest levels induced by different regulatory treatments. We also reiterate that this noteworthy result becomes apparent only via the HDTP framework, since the other two frameworks identify a different sub-model without period effects (i.e.  $M_3$ ) as the specification most suitable for inferential purposes.

### *Posterior Predictions*

Next, we compare the best sub-models chosen within each estimation framework based on their posterior predictive distributions (PPDs) for several constructs of interest. Each construct corresponds to

an explicit form of the predictive function  $g(\cdot)$  on the left hand side of (19). Our first predictive outcome of interest is the probability that chosen extraction equals the socially optimal level of two, i.e.

$pr((y | \mathbf{x}_f, \mathbf{h}_f) = 2)$ . Naturally, this latter outcome is not defined for a continuous density like the HDTN.

This is an important shortcoming compared to the other two frameworks, as this probability will be of major interest in most experimental settings.<sup>13</sup> Our second outcome of interest is the probability that harvest falls within the lowest extraction tier. For the HDTN we interpret this as  $pr((y | \mathbf{x}_f, \mathbf{h}_f) < 3)$ , and for the other two frameworks as  $pr((y | \mathbf{x}_f, \mathbf{h}_f) \in \{1, 2, 3\})$ . The third predictive construct is the expected harvest level, i.e.  $E(y | \mathbf{x}_f, \mathbf{h}_f)$ , where the expectation over harvest is to be understood as the expectation over the stochastic elements of the likelihood function, conditional on random effects  $\gamma_i$  and a given draw of model parameters  $\theta$ . The posterior predictive simulator then removes these conditionalities by estimating expected harvest for repeated draws of  $\gamma_i$  and  $\theta$ . The exact expressions for our posterior predictive constructs are given in Appendix D for each of the three frameworks.

We generate these PPDs for each regulatory treatment by setting treatment indicators to the corresponding combination of zeros and ones. We also set “education” to the sample mean for a given community. To capture gender effects we create weighted averages of separate PPDs for each value of the female indicator, using the observed proportion of females as weights. For the HDTP sub-model we set the indicator for “period” to one to place a stronger focus on the second half of a given experimental run and thus allow for learning effects. Given the settings for our demographic regressors  $\mathbf{x}_f$  we would expect our posterior outcomes of interest to lie in the vicinity of sample results.

Table 5 shows the posterior means and standard deviations for predictive outcomes for all models and regions. The first three columns of the table depict corresponding sample statistics, evaluated over the last five periods of a given game. The first block of rows gives results for the probability that harvest is equal to the social optimum. Both HOP and HDTP sub-models agree with the sample outcome in the ranking of treatment effects for the Pacific region, with  $T_2$  achieving the highest posterior probability of

inducing a socially optimal choice. For CAR and MAG the two sub-models agree with each other by assigning the highest probability of achieving the social optimum to  $T_4$ , while the sample results allocate a slightly higher probability to  $T_3$  than to  $T_4$ . Overall, these predicted point probabilities are generated with considerable posterior noise as is evident from the relatively large posterior standard deviations. This is not surprising given the relatively small number of data points upon which these estimates are based.

The second block of rows depicts sample statistics and predictive posterior results for the probability that harvest lies at or below a level of three. For this outcome, the agreement between sub-models and sample results for the ranking over treatments is complete for all three regions. Specifically,  $T_2$  receives the highest posterior probability for the Pacific region, followed by  $T_3$  and  $T_4$ , while the highest probability of a lowest-tier effort is achieved under  $T_4$  for CAR and MAG, followed by  $T_3$  and  $T_2$ .

These rankings over treatments are exactly reversed for the third predictive outcome, expected harvest. Again, sample statistics and estimated sub-models largely agree in this respect for all regions. Treatment two induces the lowest expected harvest level for PAC, while  $T_4$  leads to lowest expected extraction for CAR and MAG, as predicted by the HOP. There is only a minor deviation from this pattern for the HDTP, which predicts harvest to have a slightly lower posterior mean under  $T_3$  than  $T_4$ . In all cases the open access scenario  $T_1$  is associated with the lowest probability of inducing low harvest levels and with the highest posterior mean for expected harvest.

The shared pattern of treatment rankings for all outcomes of interest across frameworks is the first key finding flowing from our posterior predictive analysis. The second key result is the fact that the HDTN generates clearly lower posterior means for probabilities associated with the lowest effort tier for all treatments (second block of rows) compared to the HOP and HDTP, and higher posterior means for expected harvest (third block of rows). This suggests that the HDTN may not allocate sufficient density mass to lower extraction levels, regardless of treatment. Our graphical analysis of predicted extraction given below confirms this conjecture.

Posterior standard deviations are generally comparable in magnitude for the HOP and the HDTP and clearly lower for the HDTN for all outcomes of interest. However, this latter result is somewhat

misleading, as these lower predictive standard deviations are not a manifestation of superior model efficiency, but rather of the HDTN's inability to allocate appropriate probability mass to boundary outcomes. This becomes obvious upon examination of the full posterior plots for these predictive outcomes.

Figure 1 depicts predictive pdfs and cdfs for expected harvest for all three frameworks, with focus on the Pacific and Magdalena regions (top and bottom sets of graphs, respectively), given these two areas' diametrically different experience with governmental versus local intervention in resource regulation. The pdfs in Figure 1 can be interpreted as the predictive density of expected harvest for a prototypical participant, controlling for unobserved heterogeneity and parameter uncertainty. The cdf plots translate these densities in the usual fashion into probabilities of unconditional expected harvest exceeding a given threshold.<sup>14</sup>

The most striking insight flowing from the figure is the HDTN's inability to allocate sufficient probability mass to harvest levels at or near the boundaries, regardless of region or treatment. As is obvious from both the pdfs and cdfs the HDTN under-estimates probability densities for both very low and very high harvest levels. Both the HOP and the HDTP show a much broader distribution of expected harvest, with considerable mass at the boundaries for some of the treatments. These boundary effects emerge as more pronounced for the HDTP than the HOP. For example, it is evident only from the HDTP that the quota treatment  $T_2$  induces a bimodal density for expected harvest for the Magdalena region, with highest probability mass allocated to the lowest and highest harvest levels. This has significant policy relevance because it indicates that a quota plus punishment treatment in that region might induce countervailing incentives for some fishermen and pronounced disagreement amongst local resource users. This bimodality is at best hinted at by the HOP and completely absent for the HDTN.

The HDTP also illustrates more clearly than the other two frameworks that the quota plus mild punishment treatment  $T_2$  and the communication treatment  $T_4$  shift the bulk of density mass for expected extraction towards the lowest tier for PAC and MAG, respectively. Looking at the cdf for Magdalena, the

HDTN falsely suggests that  $T_4$  does not stochastically dominate  $T_2$  until harvest levels of three or four, while the HOP and HDTP indicate that stochastic dominance commences at levels in the 1.5 to 2.5 range.

Overall, we thus conclude that the HDTP provides a sharper and more complete picture than the other two frameworks with respect to the distributional implications of treatments on harvest.

Distributional impacts are somewhat more diluted but still reliable for the HOP. The HDTN, in contrast, fails to portray an accurate picture of predictive densities for all treatments. It tightly and misleadingly allocates probability mass close to the center of the permissible extraction range, missing important boundary effects induced by some of the treatments.

## V) Conclusion

This study proposes a doubly-truncated count data framework to process data from Social Dilemma Games. We compare this framework to past estimation approaches based on ordered outcomes and truncated continuous densities. We generalize all three frameworks to ex ante accommodate unobserved heterogeneity in individual decision-making. For each framework we specify an identical set of sub-models, distinguished by different conditional mean functions. We examine the performance of these models using a Bayesian estimation framework and data from recent CPR field experiments implemented in different rural communities in Colombia. Our Bayesian approach circumvents estimation hurdles that traditionally plague classical estimation via maximum likelihood and allows for a rigorous comparison of sub-models within and across estimation frameworks based on marginal likelihood scores and Bayes Factors.

Several key results flow from this analysis. All three frameworks lend strong support to the presence of *treatment-specific* unobserved heterogeneity in extraction decisions. The three frameworks also agree in the ranking of the means and variances of these random treatment effects, and on the ranking of treatment-specific predictive outcomes of interest. The count data framework receives overwhelmingly higher posterior probability than the other two frameworks as judged by marginal likelihood scores and Bayes Factors for any of the specified sub-models. This translates into more

efficient posterior results for all parameters of interest. It is also noteworthy that the best, i.e. “most likely” sub-models identified within each framework are not identical. Specifically, the count data framework alone assigns highest posterior weight to a sub-model with temporal effects for chosen extraction.

Most importantly, we find that the truncated continuous framework fails to allocate adequate probability mass to predicted harvest levels near the boundaries, which leads to an under-estimation of probabilities for both low and high extraction levels. This is an important shortcoming given that in most Social Dilemma Games the social optimum and the pure strategy Nash equilibrium are near the boundaries of the permissible action space. Consequently, predicted boundary probabilities traditionally receive considerable attention. The ordered outcome framework is not plagued by this boundary problem, but it misses other important features embedded in the data, such as the multi-modality for expected harvest for one of the field experiments. The count framework clearly generates the most informative predictive densities for the outcomes of interest examined in our analysis.

In summary, we conclude that, at least for our application, the choice of framework is of secondary importance if the analyst is primarily interested in a correct ranking of regulatory treatment effects and heterogeneity noise surrounding each effect. However, for any inference beyond this basic level a framework that is better suited to accommodate the bounded integer nature of the outcome variable will be needed. Our proposed hierarchical doubly-truncated Poisson model performs well in that it efficiently addresses the type of key research questions that traditionally motivate the original experiment. We surmise that the bulk of our results are of general relevance for data from Social Dilemma games. The application of the HDTP framework and related count data specifications to other SD data will be subject to future research.

### Appendix A: Posterior Simulator for the HDTN Framework

To draw from  $p(\boldsymbol{\beta} | \mathbf{y}^*, \mathbf{X}, \mathbf{H}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\Sigma})$  unconditional on the random coefficients one can express the basic model at the individual panel level (i.e. all  $p=1 \dots P$  observations associated with individual  $i$ ) as

$$\begin{aligned} \mathbf{y}_i^* &= \mathbf{X}_i \boldsymbol{\beta} + \mathbf{H}_i \boldsymbol{\gamma} + \mathbf{H}_i \boldsymbol{\alpha}_i + \boldsymbol{\varepsilon}_i, \quad \text{where} \\ \boldsymbol{\alpha}_i &= \boldsymbol{\gamma}_i - \boldsymbol{\gamma}, \quad \text{and} \quad \mathbf{H}_i \boldsymbol{\alpha}_i \sim mvn(\mathbf{0}, \mathbf{H}_i \boldsymbol{\Sigma} \mathbf{H}_i'). \end{aligned} \quad (\text{A1})$$

Applying standard expressions for the posterior moments of linear regression models with independent normal- $ig$  priors (e.g. Lindley and Smith, 1972), we obtain:

$$\begin{aligned} \boldsymbol{\beta} | \mathbf{y}^*, \mathbf{X}, \mathbf{H}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\Sigma} &\sim mvn(\boldsymbol{\mu}_1, \mathbf{V}_1) \quad \text{where} \\ \mathbf{V}_1 &= \left( \mathbf{V}_\beta^{-1} + \sum_{i=1}^n \mathbf{X}_i' \left( \mathbf{H}_i (\boldsymbol{\Sigma} + \sigma^2 \mathbf{I}_p) \mathbf{H}_i' \right)^{-1} \mathbf{X}_i \right)^{-1} \quad \text{and} \\ \boldsymbol{\mu}_1 &= \mathbf{V}_1 \left( \mathbf{V}_\beta^{-1} \boldsymbol{\mu}_\beta + \sum_{i=1}^n \mathbf{X}_i' \left( \mathbf{H}_i (\boldsymbol{\Sigma} + \sigma^2 \mathbf{I}_p) \mathbf{H}_i' \right)^{-1} (\mathbf{y}_i^* - \mathbf{H}_i \boldsymbol{\gamma}) \right). \end{aligned} \quad (\text{A2})$$

A similar conceptual re-writing of the baseline model allows for the derivation of the posterior distribution of  $\boldsymbol{\gamma}$ :

$$\begin{aligned} \boldsymbol{\gamma} | \mathbf{y}^*, \mathbf{X}, \mathbf{H}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\Sigma} &\sim mvn(\boldsymbol{\mu}_1, \mathbf{V}_1) \quad \text{where} \\ \mathbf{V}_1 &= \left( \mathbf{V}_\gamma^{-1} + \sum_{i=1}^n \mathbf{H}_i' \left( \mathbf{H}_i (\boldsymbol{\Sigma} + \sigma^2 \mathbf{I}_p) \mathbf{H}_i' \right)^{-1} \mathbf{H}_i \right)^{-1} \quad \text{and} \\ \boldsymbol{\mu}_1 &= \mathbf{V}_1 \left( \mathbf{V}_\gamma^{-1} \boldsymbol{\mu}_\gamma + \sum_{i=1}^n \mathbf{H}_i' \left( \mathbf{H}_i (\boldsymbol{\Sigma} + \sigma^2 \mathbf{I}_p) \mathbf{H}_i' \right)^{-1} (\mathbf{y}_i^* - \mathbf{X}_i \boldsymbol{\beta}) \right). \end{aligned} \quad (\text{A3})$$

By analogous reasoning, the conditional posterior distribution for individual-level random coefficients can be derived as

$$\begin{aligned} \boldsymbol{\gamma}_i | \mathbf{y}_i^*, \mathbf{X}_i, \mathbf{H}_i, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\Sigma} &\sim mvn(\boldsymbol{\mu}_1, \mathbf{V}_1) \quad \text{where} \\ \mathbf{V}_1 &= \left( \boldsymbol{\Sigma}^{-1} + \frac{1}{\sigma^2} \mathbf{H}_i' \mathbf{H}_i \right)^{-1} \quad \text{and} \\ \boldsymbol{\mu}_1 &= \mathbf{V}_1 \left( \boldsymbol{\Sigma}^{-1} \boldsymbol{\gamma} + \frac{1}{\sigma^2} \mathbf{H}_i' (\mathbf{y}_i^* - \mathbf{X}_i \boldsymbol{\beta}) \right). \end{aligned} \quad (\text{A4})$$

It should be noted that the hierarchical moments  $\boldsymbol{\gamma}$  and  $\boldsymbol{\Sigma}$  now play the role of the hyper-priors for  $\boldsymbol{\gamma}$  in (A3). Conditional on  $\boldsymbol{\beta}$  and  $\boldsymbol{\gamma}_i$  draws of  $\sigma^2$  can be obtained from an updated inverse-gamma density via:

$$\sigma^2 | \mathbf{y}^*, \mathbf{X}, \mathbf{H}, \boldsymbol{\beta}, \boldsymbol{\gamma}_i = ig(\eta_1, \kappa_1) \quad \text{where} \quad (A5)$$

$$\eta_1 = \frac{2\eta_0 + n}{2}, \quad \kappa_1 = \frac{2\kappa_0 + \sum_{i=1}^N (\mathbf{y}_i^* - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{H}_i \boldsymbol{\gamma}_i)' (\mathbf{y}_i^* - \mathbf{X}_i \boldsymbol{\beta} - \mathbf{H}_i \boldsymbol{\gamma}_i)}{2}$$

To draw  $\boldsymbol{\Sigma}_{jj}, j=1 \dots k_r$  we need to work with all deviations of individual-specific random coefficients from their grand mean. The conditional posterior emerges as

$$\boldsymbol{\Sigma}_{jj} | \gamma_j, \gamma_{ij} \sim ig(\nu_1, \varphi_1) \quad \text{where} \quad (A6)$$

$$\nu_1 = \frac{2\nu_0 + n}{2}, \quad \varphi_1 = \frac{2\varphi_0 + \sum_{i=1}^n (\gamma_{ij} - \gamma_j)' (\gamma_{ij} - \gamma_j)}{2}$$

The final step of a given round of the posterior simulator constitutes of draws of latent data  $\mathbf{y}^*$  for all observations that are located either at the lower or upper bound of the permissible range for harvest. A latent data point associated with an observed value of  $y_{ip} = E_{\min}$  is simply obtained through draws from the truncated normal distribution with mean  $\mathbf{x}_i' \boldsymbol{\beta} + \mathbf{h}_{ip}' \boldsymbol{\gamma}_i$ , standard deviation  $\sigma$ , and upper truncation bound  $E_{\min}$ . Analogously, latent data for cases with  $y_{ip} = E_{\max}$  are drawn from a normal density with the same moments, but with a lower truncation bound of  $E_{\max}$ .

### Appendix B: Posterior Simulator for the HOP Framework

The first six steps of the Gibbs Sampler for the HOP follow exactly those outlined for the HDTN, with all parameters and latent data replaced by their re-parameterized counterparts, and  $\sigma$  replaced by  $\delta$ .

The conditional posterior for the bin thresholds does not take a common form. A Metropolis-Hastings algorithm is needed to obtain draws from this distribution. Given the truncated normal prior as shown in (9) the conditional posterior kernel for a given threshold  $\tilde{c}_b, b = 1 \cdots B - 1$ , can be expressed as the product of the prior times the relevant part of the likelihood function, i.e.

$$p\left(\tilde{c}_b \mid \mathbf{y}, \tilde{\mathbf{y}}^*, \mathbf{X}, \mathbf{H}, \tilde{\boldsymbol{\beta}}, \delta^2, \tilde{\boldsymbol{\gamma}}_i\right) \propto \frac{\exp\left(-\frac{1}{2\tilde{v}_c^2} \tilde{c}_b^2\right)}{F(\tilde{c}_{b+1}) - F(\tilde{c}_{b-1})} \prod_{ip: y_{ip} = b+1} \Phi\left(\frac{\tilde{c}_b - (\mathbf{x}'_i \tilde{\boldsymbol{\beta}} + \mathbf{h}'_{ip} \tilde{\boldsymbol{\gamma}}_i)}{\delta}\right) - \Phi\left(\frac{\tilde{c}_{b-1} - (\mathbf{x}'_i \tilde{\boldsymbol{\beta}} + \mathbf{h}'_{ip} \tilde{\boldsymbol{\gamma}}_i)}{\delta}\right) \quad (\text{B1})$$

$$* \prod_{ip: y_{ip} = b+2} \Phi\left(\frac{\tilde{c}_{b+1} - (\mathbf{x}'_i \tilde{\boldsymbol{\beta}} + \mathbf{h}'_{ip} \tilde{\boldsymbol{\gamma}}_i)}{\delta}\right) - \Phi\left(\frac{\tilde{c}_b - (\mathbf{x}'_i \tilde{\boldsymbol{\beta}} + \mathbf{h}'_{ip} \tilde{\boldsymbol{\gamma}}_i)}{\delta}\right)$$

A random-walk type proposal function worked well for our application. We draw a candidate  $\tilde{c}_{b,c}$  from a normal proposal density  $q(\tilde{c}_{b,c} \mid \tilde{c}_b)$  with mean  $\tilde{c}_b$  (the current draw), a pre-specified standard deviation  $s_c$ , and truncation to the interval  $[\tilde{c}_{b-1}, \tilde{c}_{b+1}]$ . We set  $s_c = 0.8(1/n)$ , which yielded desirable acceptance rates of approximately 40-40% for all threshold parameters (see e.g. Gelman et al., 2004, p. 306). The

candidate draw is accepted with probability  $\alpha = \min\left\{\frac{p(\tilde{c}_{b,c} \mid \mathbf{y}, \tilde{\mathbf{y}}^*, \mathbf{X}, \mathbf{H}, \tilde{\boldsymbol{\beta}}, \delta^2, \tilde{\boldsymbol{\gamma}}_i) q(\tilde{c}_b \mid \tilde{c}_{b,c})}{p(\tilde{c}_b \mid \mathbf{y}, \tilde{\mathbf{y}}^*, \mathbf{X}, \mathbf{H}, \tilde{\boldsymbol{\beta}}, \delta^2, \tilde{\boldsymbol{\gamma}}_i) q(\tilde{c}_{b,c} \mid \tilde{c}_b)}, 1\right\}$ .

The final step of a given round of the posterior simulator constitutes of draws of latent data  $\mathbf{y}^*$ . As described in Li and Tobias (2008) a latent data point associated with an observed value of  $y_{ip} = e$  is simply obtained through draws from the truncated normal distribution with mean  $\mathbf{x}'_i \tilde{\boldsymbol{\beta}} + \mathbf{h}'_{ip} \tilde{\boldsymbol{\gamma}}_i$ , standard deviation  $\delta$ , and lower and upper truncation bounds  $\tilde{c}_{e-2}$  and  $\tilde{c}_{e-1}$ , respectively.

### Appendix C: Posterior Simulator for the HDTP Framework

The relevant part of the (un-normalized) posterior distribution for a given set of random effects  $\gamma_i$  is given by

$$p(\gamma_i | y_i, \mathbf{X}_i, \mathbf{H}_i, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}) \propto f(\gamma_i | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) \prod_{p=1}^P \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i)^{y_{ip}}}{y_{ip}! \left( \sum_{k=L}^U \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i)^k}{k!} \right)} \quad (\text{C1})$$

where  $f(\cdot)$  denotes the multivariate normal density. The ‘‘tailored proposal’’ implementation of the MH algorithm requires finding the mode of the logged form of this target density (e.g. Chib et al., 1998). This can be accomplished with a Newton-Raphson sub-routine, using the following gradient and Hessian:

$$\begin{aligned} g(\gamma_i) &= -\boldsymbol{\Sigma}^{-1}(\gamma_i - \boldsymbol{\gamma}) + \sum_{p=1}^P \left( y_{ip} - \left( \sum_{k=L}^U \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i)}{k!} \right)^{-1} \left( \sum_{k=L}^U \frac{k \exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i))}{k!} \right) \right) \mathbf{h}_{ip} \\ H(\gamma_i) &= -\boldsymbol{\Sigma}^{-1} - \sum_{p=1}^P \left( \sum_{k=L}^U \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i)}{k!} \right)^{-2} \\ &\quad \left( \left( \sum_{k=L}^U \frac{\exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i))}{k!} \right) \left( \sum_{k=L}^U \frac{k^2 \exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i))}{k!} \right) - \left( \sum_{k=L}^U \frac{k \exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i))}{k!} \right)^2 \right) \mathbf{h}_{ip} \mathbf{h}'_{ip} \end{aligned} \quad (\text{C2})$$

The resulting mode  $\hat{\gamma}_i$  is then used as the mean in a multivariate t-distribution that forms the basis for proposal draws of candidate vector  $\tilde{\gamma}_i$ . Specifically we use Chib et al.'s (1998) Acceptance-Rejection version of finding suitable candidate draws in our application.

A similar approach is taken to obtain draws of fixed effects vector  $\boldsymbol{\beta}$ . The relevant part of the un-normalized posterior distribution, conditional on  $\gamma_i$  is given as

$$p(\boldsymbol{\beta} | y_i, \mathbf{X}_i, \mathbf{H}_i, \boldsymbol{\Sigma}, \boldsymbol{\gamma}, \gamma_i) \propto f(\boldsymbol{\beta} | \boldsymbol{\mu}_\beta, \mathbf{V}_\beta) \prod_{i=1}^n \left( \prod_{p=1}^P \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i)^{y_{ip}}}{y_{ip}! \left( \sum_{k=L}^U \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \gamma_i)^k}{k!} \right)} \right) \quad (\text{C3})$$

The mode of the logged version of this density is again obtained via a Newton-Raphson sub-routine with gradient and Hessian given as

$$\begin{aligned}
g(\boldsymbol{\beta}) &= -\mathbf{V}_{\boldsymbol{\beta}}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_{\boldsymbol{\beta}}) + \sum_{i=1}^n \sum_{i=1}^n \left( y_{ip} - \left( \sum_{k=L}^U \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i)}{k!} \right)^{-1} \left( \sum_{k=L}^U \frac{k \exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i))}{k!} \right) \right) \mathbf{x}_i \\
H(\boldsymbol{\beta}) &= -\mathbf{V}_{\boldsymbol{\beta}}^{-1} - \sum_{i=1}^n \sum_{p=1}^P \left( \sum_{k=L}^U \frac{\exp(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i)}{k!} \right)^{-2} \\
&\quad \left( \left( \sum_{k=L}^U \frac{\exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i))}{k!} \right) \left( \sum_{k=L}^U \frac{k^2 \exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i))}{k!} \right) - \left( \sum_{k=L}^U \frac{k \exp(k(\mathbf{x}'_i \boldsymbol{\beta} + \mathbf{h}'_{ip} \boldsymbol{\gamma}_i))}{k!} \right)^2 \right) \mathbf{x}_i \mathbf{x}'_i
\end{aligned} \tag{C4}$$

The resulting mode  $\hat{\boldsymbol{\beta}}$  figures in a multivariate t-density from which a candidate draw  $\tilde{\boldsymbol{\beta}}$  is taken.

Specifically,  $\hat{\boldsymbol{\beta}}$  can be the mean of this proposal density in the basic version of the tailored MH implementation. Alternatively – as implemented in our application – the mean of the multivariate t density can be formed by reflecting the current value of  $\boldsymbol{\beta}$  around  $\hat{\boldsymbol{\beta}}$ . We refer the reader again to Chib et al. (1998) for further details.

The final two steps of the Gibbs Sampler are straightforward: Draws from  $p(\boldsymbol{\gamma} | \mathbf{y}, \mathbf{X}, \mathbf{H}, \boldsymbol{\Sigma}, \boldsymbol{\gamma}_i)$  can be obtained from a multivariate normal density with variance and mean given as

$$\mathbf{V}_1 = (\mathbf{V}_{\boldsymbol{\gamma}}^{-1} + n\boldsymbol{\Sigma}^{-1})^{-1} \quad \text{and} \quad \boldsymbol{\mu}_1 = \mathbf{V}_1 \left( \mathbf{V}_{\boldsymbol{\gamma}}^{-1} \boldsymbol{\mu}_{\boldsymbol{\gamma}} + \boldsymbol{\Sigma}^{-1} \sum_{i=1}^n \boldsymbol{\gamma}_i \right), \tag{C5}$$

and draws of variance terms contained in  $\boldsymbol{\Sigma}$  follow the same procedure as outlined for the HDTN in Appendix A.

### Appendix D: Posterior Predictive Constructs and Implementation

*HDTN:*

The probability that harvest does not exceed a level of three, conditional on  $\gamma_i$  and  $\boldsymbol{\theta}$ , is given as:

$$pr\left((y | \mathbf{x}_f, \mathbf{h}_f, \gamma_i, \boldsymbol{\theta}) \leq 3; 1 \leq y \leq 9\right) = \frac{\Phi\left(\frac{3-\omega}{\sigma}\right) - \Phi\left(\frac{1-\omega}{\sigma}\right)}{\Phi\left(\frac{9-\omega}{\sigma}\right) - \Phi\left(\frac{1-\omega}{\sigma}\right)} \quad \text{where} \quad (D1)$$

$$\omega = \mathbf{x}'_f \boldsymbol{\beta} + \mathbf{h}'_f \gamma_i$$

The expression for expected harvest is:

$$E(y | \mathbf{x}_f, \mathbf{h}_f, \gamma_i, \boldsymbol{\theta}; 1 \leq y \leq 9) = \omega + \sigma \frac{\phi\left(\frac{1-\omega}{\sigma}\right) - \phi\left(\frac{9-\omega}{\sigma}\right)}{\Phi\left(\frac{9-\omega}{\sigma}\right) - \Phi\left(\frac{1-\omega}{\sigma}\right)} \quad (D2)$$

*HOP:*

The probability that harvest equals two conditional on  $\gamma_i$  and  $\boldsymbol{\theta}$ , is given as:

$$pr\left((y | \mathbf{x}_f, \mathbf{h}_f, \gamma_i, \boldsymbol{\theta}) = 2\right) = \Phi(c_1 - \omega) - \Phi(-\omega) \quad (D3)$$

where  $c_1$  is the first estimated threshold in original parameterization and  $\omega$  is defined as above.

Similarly, the probability of harvest falling into the lowest tier can be expressed as:

$$pr\left((y | \mathbf{x}_f, \mathbf{h}_f, \gamma_i, \boldsymbol{\theta}) \in \{1, 2, 3\}\right) = \Phi(c_2 - \omega) \quad (D4)$$

Expected harvest is given as:

$$E(y | \mathbf{x}_f, \mathbf{h}_f, \gamma_i, \boldsymbol{\theta}) = \sum_{e=1}^9 e \cdot pr(y = e) \quad (D5)$$

where the bin-probabilities  $pr(y = e)$  are given in the main text, equation (5).

*HDTP:*

For the HDTP, our three constructs of interest take the following form:

$$\begin{aligned}
pr((y | \mathbf{x}_f, \mathbf{h}_f, \boldsymbol{\gamma}_i, \boldsymbol{\theta}) = 2; 1 \leq y \leq 9) &= \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\lambda^k}{k!} \right)^{-1} \frac{\lambda^2}{2!} \\
pr((y | \mathbf{x}_f, \mathbf{h}_f, \boldsymbol{\gamma}_i, \boldsymbol{\theta}) \in \{1, 2, 3\}; 1 \leq y \leq 9) &= \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\lambda^k}{k!} \right)^{-1} \sum_{j=1}^3 \frac{\lambda^j}{j!} \\
E(y | \mathbf{x}_f, \mathbf{h}_f, \boldsymbol{\gamma}_i, \boldsymbol{\theta}; 1 \leq y \leq 9) &= \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\lambda^k}{k!} \right)^{-1} \sum_{y=E_{\min}}^{E_{\max}} \frac{y \lambda^y}{y!}
\end{aligned} \tag{D6}$$

where  $\lambda = \exp(\mathbf{x}'_f \boldsymbol{\beta} + \mathbf{h}'_f \boldsymbol{\gamma}_i)$

*Implementation:*

In practice the required marginalization over both heterogeneous coefficients  $\boldsymbol{\gamma}_i$  and parameters  $\boldsymbol{\theta}$  can be accomplished as follows for any conditional construct of interest and framework, i.e. for any form of  $g(\mathbf{x}_f, \mathbf{h}_f | \boldsymbol{\gamma}_i, \boldsymbol{\theta})$ :

Step 1: For each set of parameter draws  $\boldsymbol{\theta}_r$  flowing from the main posterior simulator draw  $r_2$  terms of  $\boldsymbol{\gamma}_i$  from  $mvn(\boldsymbol{\gamma}_r, \boldsymbol{\Sigma}_r)$ , where subscript  $r$  indicates the  $r^{\text{th}}$  draw in the original Gibbs Sampler.<sup>15</sup>

Step 2: For each draw of  $\boldsymbol{\gamma}_i$  compute  $g(\mathbf{x}_f, \mathbf{h}_f | \boldsymbol{\gamma}_i, \boldsymbol{\theta})$ .

Step 3: Repeat these two steps for each of the  $r=1..R$  original draws of parameters.

This yields  $R * r_2$  draws of probabilities for the effort tier of interest. This sequence can then be examined with respect to shape and moments, analogous to the predictive posteriors of the original model parameters.

## Appendix E: Payoff Table for the Colombia CPR Experiment

Total extraction by others	My level of extraction									Average of the others
	1	2	3	4	5	6	7	8	9	
4	900	996	1087	1172	1252	1326	1395	1458	1516	1.0
5	882	976	1064	1146	1223	1295	1361	1421	1476	1.3
6	864	955	1040	1120	1194	1263	1326	1384	1436	1.5
7	846	934	1017	1094	1165	1231	1292	1347	1396	1.8
8	829	914	994	1068	1137	1200	1258	1310	1357	2.0
9	811	893	970	1042	1108	1168	1223	1273	1317	2.3
10	793	873	947	1016	1079	1137	1189	1236	1277	2.5
11	775	852	923	989	1050	1105	1154	1198	1237	2.8
12	757	831	900	963	1021	1073	1120	1161	1197	3.0
13	739	811	877	937	992	1042	1086	1124	1157	3.3
14	721	790	853	911	963	1010	1051	1087	1117	3.5
15	703	769	830	885	934	978	1017	1050	1077	3.8
16	686	749	807	859	906	947	983	1013	1038	4.0
17	668	728	783	833	877	915	948	976	998	4.3
18	650	708	760	807	848	884	914	939	958	4.5
19	632	687	736	780	819	852	879	901	918	4.8
20	614	666	713	754	790	820	845	864	878	5.0
21	596	646	690	728	761	789	811	827	838	5.3
22	578	625	666	702	732	757	776	790	798	5.5
23	560	604	643	676	703	725	742	753	758	5.8
24	543	584	620	650	675	694	708	716	719	6.0
25	525	563	596	624	646	662	673	679	679	6.3
26	507	543	573	598	617	631	639	642	639	6.5
27	489	522	549	571	588	599	604	604	599	6.8
28	471	501	526	545	559	567	570	567	559	7.0
29	453	481	503	519	530	536	536	530	519	7.3
30	435	460	479	493	501	504	501	493	479	7.5
31	417	439	456	467	472	472	467	456	439	7.8
32	400	419	433	441	444	441	433	419	400	8.0
33	382	398	409	415	415	409	398	382	360	8.3
34	364	378	386	389	386	378	364	345	320	8.5
35	346	357	362	362	357	346	329	307	280	8.8
36	328	336	339	336	328	314	295	270	240	9.0

The pure strategy static Nash Equilibrium = 7. The joint-payoff maximizing level = 2.

## Notes

<sup>1</sup> For a semi-parametric approach to estimate hierarchical integer models see e.g. Gurmu et al. (1999) and Jochmann and León-González (2004).

<sup>2</sup> In some CPR games the specified integer range denotes extraction *effort*, for example in terms of temporal units per season or year, as opposed to actual harvest levels. This does not affect the general econometric concerns raised in this study. For consistency with our empirical application we will use the “harvest” or “extraction” interpretation of decision levels throughout the remainder of this text.

<sup>3</sup> The covariances in  $\Sigma$  are set to zero for our application since each individual participates in only two of four treatments. Thus, different covariance terms would apply to different sub-samples of the data, which would make their interpretation problematic. In theory, a separate within-group variance matrix could be estimated for each group in a given experimental session. However, in our case this would lead to an intractable proliferation of poorly identified parameters given the large number of up to 12 groups per treatment and the small number of five players per group. Also, it would be straightforward to add a second hierarchical layer if there is additional nesting in the data. For example, if the same experiment is conducted in several communities, one could specify group means of heterogeneous effects to be drawn from a community-level distribution. In our application the number of communities (three) is too small to allow for the efficient estimation of second-layer parameters. Instead, we focus on an informal comparison of community-level results.

<sup>4</sup> As discussed in Chib and Carlin (1999) in the context of related hierarchical regression models, this further improves the mixing of individual draws and thus the convergence speed of the Gibbs Sampler. A detailed discussion of marginalizing draws of fixed regression parameters or parameters of higher hierarchical levels over lower level random coefficients is given in Lindley and Smith (1972).

<sup>5</sup> For a comparative review of alternative approaches to approximate the marginal likelihood see Han and Carlin (2001).

<sup>6</sup> For comparison, a day's wage in the three regions at the time of the experiments varied between 10,000 and 15,000 pesos.

<sup>7</sup> From a game-theoretic perspective the expected penalties under regulatory treatments  $T_2$  and  $T_3$  are not high enough to induce compliance with the harvest quota. However, such weak enforcement strategies are characteristic of regulatory control of natural resources in the developing world.

<sup>8</sup> We also included subjects' age in preliminary specification but found no measurable effects for this regressor.

<sup>9</sup> Posterior summaries for the threshold coefficients in the HOP and the error variance for the HDTN are omitted for ease of exposition. They are available from the authors upon request.

<sup>10</sup> Naturally, the numerical values of these effects will differ between the HDTN and the HOP given the necessary restriction of  $\sigma^2 = 1$  for the latter framework.

<sup>11</sup> One minor exception is Magdalena for the HDTP, where the ranking between  $T_3$  and  $T_4$  is reversed compared to the other two regions.

<sup>12</sup> As shown in List and Price (Forthcoming) the role of trust-building to reach socially desirable outcomes in SD settings is especially important when social connections in a community are relatively weak and / or when "outsiders" from different cultural or socio-economic segments of the population participate in decision-making for a PG or a CPR.

<sup>13</sup> One could replace this exact probability with a cumulative density bounded between  $2 - \varepsilon$  and  $2 + \varepsilon$ . However, this would entail the arbitrary choice of  $\varepsilon$ . We thus abstract from deriving this probability for the HDTN.

<sup>14</sup> We use an Epanechnikov kernel density estimator with Silverman's (1986) rule-of-thumb bandwidth selector to generate these plots. We are grateful to Justin Tobias for providing Matlab code on his course web site to perform this kernel density estimation.

<sup>15</sup> In theory, only one draw of  $\gamma_i$  is required for each  $\theta_r$  to simulate the desired predictive distribution of bin probabilities. However, multiple draws are computationally inexpensive and improve the accuracy of forecasts.

References:

- Albert, J. and S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association* **88**: 669-679.
- Anderson, S. P., J. K. Goeree and C. A. Holt. 1998. A theoretical analysis of altruism and decision error in public goods games. *Journal of Public Economics* **70**: 292-323.
- Andreoni, J. 1993. An experimental test of the public-goods crowding-out hypothesis. *American Economic Review* **83**: 1317-27.
- Andreoni, J. 1995. Cooperation in public-goods experiments: Kindness or confusion? *American Economic Review* **85**: 891-904.
- Bardsley, N. and P. G. Moffatt. 2007. The experimetrics of public goods: Inferring motivations from contributions. *Theory and Decisions* **62**: 161-193.
- Brandts, J. and A. Schram. 2001. Cooperation and noise in public goods experiments: Applying the contribution function approach. *Journal of Public Economics* **79**: 399-427.
- Cameron, A. C. and P. K. Trivedi. 1998. *Regression Analysis of Count Data*. Cambridge University Press:
- Cardenas, J.-C., T. K. Ahn and E. Ostrom. 2004. Communication and cooperation in a common-pool resource dilemma: A field experiment. In *Advances in Understanding Strategic Behaviour: Game Theory, Experiments, and Bounded Rationality: Essays in Honour of Werner Gürth*, S. Huck (ed). Palgrave: New York.
- Cardenas, J. C., J. Stranlund and C. Willis. 2000. Local environmental control and institutional crowding-out. *World Development* **28**: 1719-1733.
- Cardenas, J. C., J. Stranlund and C. Willis. 2002. Economic inequality and burden-sharing in the provision of local environmental quality. *Ecological Economics* **40**: 379-395.
- Carpenter, J. 2004. When in Rome: Conformity and the provision of public goods. *Journal of Socio-Economics* **33**: 395-408.
- Casari, M. and C. P. Plott. 2003. Decentralized management of common property resources: Experiments with a centuries-old institution. *Journal of Economic Behavior and Organization* **51**: 217-247.
- Chib, S. 1995. Marginal likelihood from the Gibbs Sampler. *Journal of the American Statistical Association* **90**: 1313-1321.
- Chib, S. and B. P. Carlin. 1999. On MCMC sampling in hierarchical longitudinal models. *Statistics and Computing* **9**: 17-26.
- Chib, S., E. Greenberg and R. Winkelmann. 1998. Posterior simulation and Bayes Factors in panel count data models. *Journal of Econometrics* **86**: 33-54.
- Chib, S. and I. Jeliazkov. 2001. Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association* **95**: 270-281.

- Fehr, E. and S. Gächter. 2000. Cooperation and punishment in public goods experiments. *American Economic Review* **90**: 980-994.
- Ferraro, P. and C. Vossler (2007). Stylized facts and identification in public goods experiments: The confusion confound, Working paper, Department of Economics, University of Tennessee, Knoxville.
- Fischbacher, U., S. Gächter and E. Fehr. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* **71**: 397-404.
- Gelman, A., J. B. Carlin, H. S. Stern and D. B. Rubin. 2004. *Bayesian Data Analysis*. Chapman & Hall/CRC: Boca Raton, London, New York, Washington, D.C..
- Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics 4*, J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (eds). Oxford University Press: Oxford, UK.
- Geweke, J. 2005. *Contemporary Bayesian Econometrics and Statistics*. Wiley Interscience.
- Gurmu, S., P. Rilstone and S. Stern. 1999. Semiparametric estimation of count regression models. *Journal of Econometrics* **88**: 123-150.
- Han, C. and B. P. Carlin. 2001. Markov Chain Monte Carlo methods for computing Bayes Factors: A comparative review. *Journal of the American Statistical Association* **96**: 1122-1132.
- Harrison, G. W. and J. A. List. 2004. Field experiments. *Journal of Economic Literature* **42**: 1009-1055.
- Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, N. S. Henrich, K. Hill, F. Gil-White, M. Gurven, F. W. Marlowe, J. Q. Patton and D. Tracer. 2005. "Economic Man" In cross-cultural perspective: Behavioral experiments in 15 small-scale societies. *Behavioral and Brain Sciences* **28**: 795-855.
- Huang, H.-C. and S.-C. Lin. 2006. Time-varying discrete monetary policy reaction functions. *Applied Economics* **38**: 449-464.
- Jochmann, M. and R. León-González. 2004. Estimating the demand for health care with panel data: A semiparametric Bayesian approach. *Health Economics* **13**: 1003-1014.
- Kass, R. E. and A. E. Raftery. 1995. Bayes factors. *Journal of the American Statistical Association* **90**: 773-795.
- Koop, G. 2004. *Bayesian Econometrics*. John Wiley & Sons, Ltd:
- Koop, G., D. J. Poirier and J. L. Tobias. 2007. *Bayesian Econometric Methods*. Cambridge University Press:
- Kurzban, R., K. McCabe, V. L. Smith and B. J. Wilson. 2001. Incremental commitment and reciprocity in a real-time public goods game. *Personality and Social Psychology Bulletin* **27**: 1662-1673.
- Lancaster, T. 2004. *An Introduction to Modern Bayesian Econometrics*. Blackwell Publishing:

- Ledyard, J. O. 1995. Public goods: A survey of experimental research. In *Handbook of Experimental Economics*, J. Kagel and A. Roth (eds.). Princeton University Press:
- Li, M. and J. L. Tobias. 2006. Bayesian analysis of structural effects in an ordered equation system. *Studies in Nonlinear Dynamics & Econometrics* **10**: 1-22.
- Li, M. and J. L. Tobias. 2008. Bayesian analysis of treatment effects in an ordered potential outcome model. In *Advances in Econometrics: Modeling and Evaluating Treatment Effects in Econometrics*, D. Millimet, J. Smith and E. Vytlacil (eds).
- Lindley, D. V. and A. F. M. Smith. 1972. Bayes estimates for the linear model. *Journal of the Royal Statistical Society, Series B (Methodological)* **34**: 1-41.
- List, J. A. 2006. Field experiments: A bridge between lab and naturally occurring data. *B.E. Journal of Economic Analysis & Policy* **6**: 1-45.
- List, J. A. and M. K. Price. Forthcoming. The role of social connections in charitable fundraising: Evidence from a natural field experiment. *Journal of Economic Behavior and Organization*.
- Messer, K. D., T. M. Schmit and H. M. Kaiser. 2005. Optimal institutional mechanisms for funding generic advertising: An experimental analysis. *American Journal of Agricultural Economics* **87**: 1046-1060.
- Nandram, B. and M.-H. Chen. 1996. Reparameterizing the generalized linear model to accelerate Gibbs Sampler convergence. *Journal of Statistical Computation and Simulation* **54**: 129-144.
- Offerman, T., A. Schram and J. Sonnemans. 1998. Quantal response models in step-level public good games. *European Journal of Political Economy* **14**: 89-100.
- Ostmann, A. 1998. External control may destroy the commons. *Rationality and Society* **10**: 103-122.
- Palfrey, T. R. and J. E. Prisbrey. 1996. Altruism, reputation and noise in linear public goods experiments. *Journal of Public Economics* **61**: 409-427.
- Palfrey, T. R. and J. E. Prisbrey. 1997. Anomalous behavior in public goods experiments: How much and why? *American Economic Review* **87**: 829-846.
- Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.
- Tanner, M. A. and W. H. Wong. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82**: 528-550.
- Velez, M. A., J. J. Murphy and J. K. Stranlund. Forthcoming (a). Centralized and decentralized management of local common pool resources in the developing world: Experimental evidence from fishing communities in Colombia. *Economic Inquiry*.
- Velez, M. A., J. K. Stranlund and J. J. Murphy. Forthcoming (b). What motivates common pool resource users? Experimental evidence from the field. *Journal of Economic Behavior and Organization*.

Vyrastekova, J. and D. v. Soest. 2003. Centralized common-pool management and local community participation. *Land Economics* **79**: 500-514.

Winkelmann, R. 2003. *Econometric Analysis of Count Data*. Springer: Heidelberg and New York.

Table 1: Sample Statistics

	Pacific		Caribbean		Magdalena	
	mean / percent	std	mean	std	mean	std
age	43.2	11.6	35.7	13.4	42.7	15.4
years of formal education	4.1	2.3	6.2	3.6	4.3	2.9
female	10.0%	-	43.0%		20.0%	
lived in community ten years or longer	95.0%	-	78.0%	-	93.0%	-
fishing is the main household activity	94.0%	-	69.0%	-	87.0%	-
extraction , periods 1-5	5.05	2.40	5.00	2.50	5.00	2.50
extraction, periods 6-10	4.90	2.60	5.20	2.60	4.90	2.70

Table 2: Sub-Model Specifications and Sub-Model Comparison for all Frameworks and Areas

HDTN						
Model	# of param.	fixed effects	random effects	log mLH (BF*)		
				Pac	Car	Mag
M1	5	gender, edu	constant	-2710.71 (156.35)	-2623.05 (91.45)	-2706.86 (112.12)
M2	7	gender, edu, T1 - T4	-	-2635.46 (81.10)	-2698.83 (167.23)	-2690.13 (95.39)
M3	11	gender, edu	T1 - T4	-2554.37 (0.00)	-2531.60 (0.00)	-2595.06 (0.32)
M4	15	gender, edu, treat.*period	T1 - T4	-2563.05 (8.69)	-2542.03 (10.42)	-2603.78 (9.04)
M5	13	gender, edu	T1-T4, period	-2558.54 (4.18)	-2532.79 (1.19)	-2594.74 (0.00)
HOP						
Model	# of param.	fixed effects	random effects	log mLH (BF*)		
				Pac	Car	Mag
M1	11	gender, edu	constant	-2630.83 (143.04)	-2569.51 (75.47)	-2613.53 (107.32)
M2	13	gender, edu, T1 - T4	-	-2551.01 (63.22)	-2644.97 (150.92)	-2588.13 (81.92)
M3	17	gender, edu	T1 - T4	-2487.80 (0.00)	-2494.05 (0.00)	-2506.20 (0.00)
M4	21	gender, edu, treat.*period	T1 - T4	-2505.26 (17.47)	-2512.88 (18.83)	-2524.12 (17.91)
M5	19	gender, edu	T1-T4, period	-2509.83 (22.04)	-2509.62 (15.57)	-2521.65 (15.44)
HDTP						
Model		fixed effects	random effects	Log mLH (BF*)		
				Pac	Car	Mag
M1	4	gender, edu	constant	-2696.75 (667.19)	-2416.73 (400.49)	-2757.45 (671.31)
M2	6	gender, edu, T1 - T4	-	-2569.63 (540.07)	-2689.14 (672.90)	-2737.76 (651.62)
M3	10	gender, edu	T1 - T4	-2120.73 (91.17)	-2117.91 (101.67)	-2196.51 (110.37)
M4	14	gender, edu, treat.*period	T1 - T4	-2125.75 (96.19)	-2126.55 (110.32)	-2197.37 (111.22)
M5	12	gender, edu	T1-T4, period	-2029.56 (0.00)	-2016.24 (0.00)	-2086.14 (0.00)

mLH = marginal likelihood / \*BF = Bayes Factor. Based on the difference of the log-marginal likelihood of the most likely sub-model to the log-marginal likelihood of any other sub-model

Table 3: Sub-Model Comparison Across Frameworks

		# param.	logP	logLH	log(mLH)	BF*
<b>Pacific</b>						
Framework	Model					
HDTN	M3	11	-32.31	-2523.73	-2554.365	524.808
HOP	M3	17	-29.24	-2405.20	-2487.795	458.238
HDTP	M5	12	-31.74	-2001.91	-2029.557	0.000
<b>Caribbean</b>						
Framework	Model					
HDTN	M3	11	-35.17	-2500.52	-2531.601	515.363
HOP	M3	17	-29.85	-2411.94	-2494.048	477.810
HDTP	M5	12	-30.67	-1972.01	-2016.238	0.000
<b>Magdalena</b>						
Framework	Model					
HDTN	M5	13	-39.08	-2559.29	-2594.738	508.594
HOP	M3	17	-27.61	-2427.52	-2506.203	420.059
HDTP	M5	12	-30.22	-2041.69	-2086.144	0.000

P = prior, LH = sample likelihood, mLH = marginal likelihood, BF = Bayes Factor. \*Based on the difference of the log-marginal likelihood of the most likely model to the log-marginal likelihood of any other model within each region.

Table 4: Estimation Results for Most Likely Sub-Models, all Regions

	HDTN / Model 3			HOP / Model 3			HDTP / Model 5		
	Pacific	Caribb.	Magd.	Pacific	Caribb.	Magd.	Pacific	Caribb.	Magd.
	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)
<u>Fixed Effects</u>									
female	0.891 (0.441)	0.182 (0.315)	0.947 (0.363)	0.469 (0.288)	0.087 (0.192)	0.359 (0.193)	0.260 (0.140)	0.066 (0.097)	0.243 (0.109)
education	-0.921 (0.548)	0.532 (0.426)	0.187 (0.500)	-0.409 (0.367)	0.232 (0.262)	0.108 (0.271)	-0.207 (0.174)	0.148 (0.134)	-0.027 (0.147)
<u>RE Means</u>									
T1	6.375 (0.284)	5.446 (0.324)	5.517 (0.281)	2.568 (0.200)	2.031 (0.207)	1.860 (0.162)	1.992 (0.088)	1.691 (0.102)	1.750 (0.082)
T2	3.235 (0.501)	4.814 (0.781)	4.160 (0.740)	1.114 (0.339)	1.737 (0.418)	1.309 (0.340)	0.971 (0.202)	1.510 (0.269)	1.322 (0.252)
T3	4.058 (0.530)	4.204 (0.718)	3.227 (0.537)	1.496 (0.344)	1.442 (0.400)	0.937 (0.291)	1.286 (0.188)	1.272 (0.242)	1.049 (0.180)
T4	5.170 (0.378)	3.315 (0.393)	2.934 (0.530)	2.038 (0.297)	1.098 (0.296)	0.817 (0.298)	1.715 (0.127)	1.130 (0.145)	1.098 (0.152)
period	- -	- -	- -	- -	- -	- -	-0.043 (0.050)	0.079 (0.054)	0.068 (0.058)
<u>RE Variances</u>									
T1	1.053 (0.317)	1.572 (0.414)	0.807 (0.309)	0.518 (0.112)	0.582 (0.126)	0.354 (0.078)	0.094 (0.024)	0.131 (0.034)	0.095 (0.024)
T2	3.371 (1.451)	10.182 (4.207)	9.464 (3.697)	1.642 (0.612)	2.761 (1.086)	1.924 (0.745)	0.612 (0.250)	1.236 (0.592)	1.140 (0.472)
T3	3.995 (1.591)	8.039 (3.094)	4.080 (1.772)	1.723 (0.634)	2.341 (0.880)	1.257 (0.468)	0.529 (0.214)	0.901 (0.352)	0.503 (0.231)
T4	1.103 (0.577)	0.998 (0.546)	3.179 (1.449)	1.067 (0.381)	1.007 (0.369)	1.141 (0.417)	0.166 (0.062)	0.192 (0.084)	0.282 (0.116)
period	- -	- -	- -	- -	- -	- -	0.081 (0.022)	0.102 (0.028)	0.126 (0.033)

RE = random effects

mean = posterior mean / (std) = posterior standard deviation

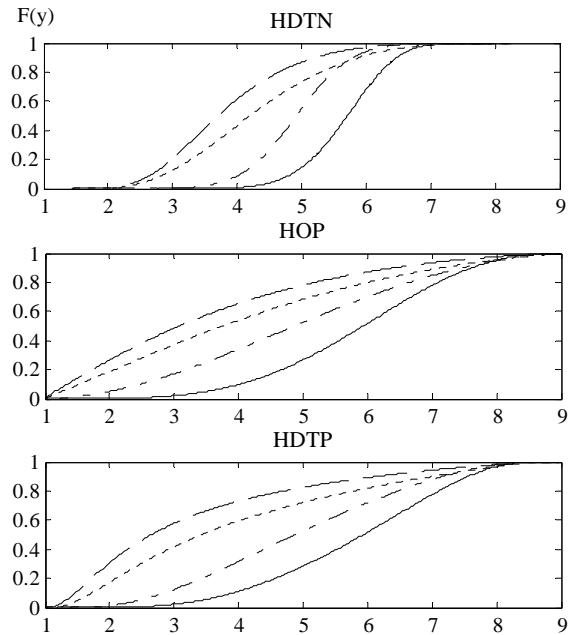
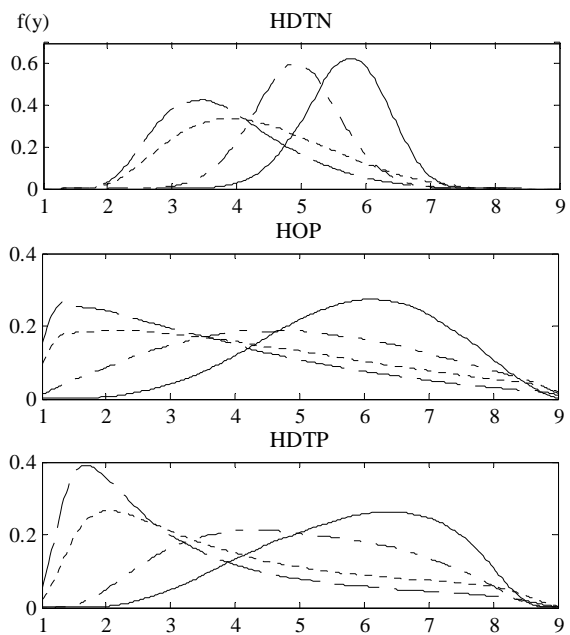
Table 5: Observed and Predicted Extraction Probabilities and Levels

	Sample			HDTN / M3			HOP / M3			HDTP / M5		
	Pacific	Caribb.	Magd.	Pacific	Caribb.	Magd.	Pacific	Caribb.	Magd.	Pacific	Caribb.	Magd.
				mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)	mean (std)
<u>Level 2</u>												
T1	0.063	0.063	0.073	N/A	N/A	N/A	0.102 (0.094)	0.090 (0.081)	0.125 (0.083)	0.061 (0.067)	0.066 (0.076)	0.071 (0.077)
T2	0.360	0.120	0.240	N/A	N/A	N/A	0.235 (0.128)	0.117 (0.110)	0.166 (0.117)	0.212 (0.106)	0.116 (0.121)	0.143 (0.123)
T3	0.320	0.240	0.270	N/A	N/A	N/A	0.211 (0.134)	0.135 (0.110)	0.205 (0.106)	0.180 (0.115)	0.140 (0.122)	0.184 (0.114)
T4	0.190	0.180	0.250	N/A	N/A	N/A	0.173 (0.130)	0.178 (0.100)	0.215 (0.102)	0.113 (0.096)	0.166 (0.106)	0.186 (0.109)
<u>Level 1-3</u>												
T1	0.226	0.183	0.243	0.101 (0.064)	0.123 (0.085)	0.143 (0.055)	0.204 (0.176)	0.223 (0.192)	0.278 (0.176)	0.183 (0.180)	0.196 (0.204)	0.210 (0.206)
T2	0.620	0.320	0.420	0.397 (0.193)	0.234 (0.234)	0.275 (0.209)	0.600 (0.318)	0.362 (0.345)	0.464 (0.335)	0.647 (0.321)	0.355 (0.366)	0.443 (0.377)
T3	0.500	0.430	0.480	0.316 (0.197)	0.264 (0.227)	0.324 (0.163)	0.509 (0.328)	0.417 (0.345)	0.558 (0.302)	0.526 (0.334)	0.422 (0.364)	0.541 (0.333)
T4	0.360	0.460	0.630	0.191 (0.098)	0.302 (0.110)	0.346 (0.149)	0.361 (0.279)	0.484 (0.289)	0.589 (0.292)	0.320 (0.252)	0.465 (0.283)	0.528 (0.304)
<u>Expected harvest</u>												
T1	5.770	5.733	5.553	5.681 (0.632)	5.520 (0.756)	5.404 (0.465)	5.862 (1.355)	5.664 (1.437)	5.568 (1.275)	5.820 (1.359)	5.795 (1.511)	5.680 (1.490)
T2	3.350	5.270	4.670	3.841 (0.980)	5.087 (1.605)	4.752 (1.354)	3.502 (1.909)	5.121 (2.447)	4.585 (2.310)	3.233 (1.800)	5.215 (2.511)	4.593 (2.462)
T3	4.170	4.900	4.280	4.301 (1.129)	4.787 (1.469)	4.291 (0.944)	4.061 (2.055)	4.672 (2.330)	3.882 (1.941)	3.904 (1.948)	4.671 (2.358)	3.818 (1.934)
T4	4.520	4.090	3.680	4.923 (0.682)	4.225 (0.611)	4.138 (0.829)	4.915 (1.836)	4.048 (1.727)	3.678 (1.840)	4.941 (1.586)	4.109 (1.585)	3.821 (1.695)

mean = posterior mean, (std) = posterior standard deviation

Figure 1: Predictive pdf and cdf for Expected Harvest

Pacific



Magdalena

