

# Institutional Heterogeneity in Social Dilemma Games: A Bayesian Examination

**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

*Draft Dec. 8, 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

# Institutional Heterogeneity in Social Dilemma Games: A Bayesian Examination

## Abstract:

A main research focus in many Social Dilemma Games is the suitability of external institutional treatments in inducing socially optimal outcomes. It is likely that participating subjects exhibit unobserved heterogeneity in their reaction to these treatments. This type of “institutional heterogeneity” has to date not found much attention in the experimental literature. We propose a Hierarchical Bayesian estimation framework to highlight these heterogeneity effects. We illustrate that models that ignore treatment-specific heterogeneity can severely under-estimate the variability in treatment-induced decisions amongst the subject population. The resulting misleading picture of comparative treatment effects can lead to sub-optimal institutional choices and related policy decisions.

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

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

## I) Introduction

Social Dilemma (SD) Experiments, such as Public Good (PG) games and Common Pool Resource (CPR) games, are designed to create a gap between the game-theoretic equilibrium strategy and the socially optimal course of action. This cleft in payoffs and the group structure of the experimental setup leaves considerable room for “other-regarding” and “other-dependent” forces such as altruism, inequity aversion, reputation-building, compliance with norms, and conditional cooperation to take root and affect observed outcomes. It is thus not surprising that in most existing applications experimentalists have focused on the identification of these different behavioral motives amongst SD game participants (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)).

However, in some SD games an equally important if not primary experimental focus rests on the effect of public institutions and policies on players’ allocation or extraction decisions (e.g. Ostrom et al., 1994, Ostmann, 1998, Beckenkamp and Ostmann, 1999, Cardenas et al., 2000, Vyrastekova and Soest, 2003, Bischoff, 2007, Velez et al., Forthcoming(a)). Such institutions are represented via exogenous treatments added to the game structure, such as quotas, penalties, open communication, and voting mechanisms. In many instances SD games with focus on policy effects are 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). Thus, field experiments with an institutional treatment component have ex ante considerable potential to inform real-world policy prescriptions for the management of the natural resource targeted by the experiment (Cardenas and Ostrom, 2004).

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. 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, Palfrey and Prisbrey, 1997, Anderson et al., 1998, Casari and Plott, 2003). In SD games with an external policy components additional heterogeneity effects are likely introduced due to the difference in how subjects react to prescribed treatments. Somewhat surprisingly, this type of “institutional heterogeneity” has to date not received much attention in the existing literature.

In this chapter we propose a fully parametric econometric framework for the analysis of SD data that is well-suited to capture unobserved subject heterogeneity in treatment effects. Using the hierarchical doubly-truncated Poisson (HDTP) model recently described in Moeltner et al. (2008) and data from a CPR field experiment in two artisanal fishing regions in Colombia we find clear evidence that subjects react to institutional treatments in heterogeneous fashion. Equally important, the extent of the resulting variability in induced decisions varies over subject populations and treatments. We also find that a “naïve” model that only allows for unobserved heterogeneity of a general nature severely under-estimates the variability in treatment-induced decisions amongst the subject population. This produces a misleading picture of comparative treatment effects on resource extraction, and could lead to sub-optimal institutional choices if the results of the SD experiment were used to inform resource policies.

We use a Bayesian estimation approach to derive these findings. The Bayesian framework offers the following analytical advantages over a classical approach: (i) it avoids the computational difficulties that would arise in Maximum Likelihood (ML) estimation of a hierarchical, nonlinear model with limited outcome space such as the HDTP, (ii) it allows for a rigorous and straightforward comparison of non-nested models via marginal likelihoods and Bayes Factors, (iii) it offers the option of using the first few rounds of a given game to generate informed priors that can then be refined using the remaining data, (iv) it can easily generate posterior distributions for predictive outcomes of interest as a simple extension of its core algorithm, and (v) it offers a rigorous statistical framework for combining predictive densities across subject populations.

The remainder of this manuscript is structured as follows: The next section introduces the econometric framework and competing models. Section III briefly describes the field experiment,

illustrates the empirical implementation of our models, and discusses estimation results. Concluding remarks are given in Section IV.

### III) Econometric Framework and Models

#### *General Framework*

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.

$y_{ip} \in \{E_{\min}, E_{\min} + 1, \dots, E_{\max} - 1, E_{\max}\}$ . 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.

Moeltner et al. (2008) show that the HDTP framework is better suited to process the data from such an experimental setup than models based on underlying continuous latent variables (e.g. Palfrey and Prisbrey, 1996, Carpenter, 2004, Bardsley and Moffatt, 2007, Velez et al., Forthcoming(a), Velez et al., Forthcoming(b)). In the HDTP, the density function for observed effort level  $y_{ip}$  is given as

$$f(y_{ip} | \lambda_{ip}, E_{\min} \leq y_{ip} \leq E_{\max}) = \frac{\exp(-\lambda_{ip}) \lambda_{ip}^{y_{ip}}}{y_{ip}! \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\exp(-\lambda_{ip}) \lambda_{ip}^k}{k!} \right)} = \frac{\lambda_{ip}^{y_{ip}}}{y_{ip}! \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\lambda_{ip}^k}{k!} \right)} \quad \text{where} \quad (1)$$

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

As shown in the last line of (1) subject-specific characteristics and treatment effects are introduced into this model via the conditional mean or “link” function  $\lambda_{ip}$ . Specifically,  $\mathbf{x}_i$  is a vector of period-invariant individual characteristics, and  $\mathbf{h}_{ip} = [T_{1,ip} \quad T_{2,ip} \quad \dots \quad T_{T,ip}]'$  is a vector of 0/1 treatment indicators that may change over periods and subjects. To be specific, for each treatment or treatment combination that

applies to person  $i$  in period  $p$ , the corresponding indicator terms in  $\mathbf{h}_{ip}$  are set to one, with the other indicators held at zero. Vector  $\boldsymbol{\beta}$  includes all fixed coefficients, and  $\boldsymbol{\gamma}_i$  comprises individual-specific random coefficients. Given the focus of this study these random coefficients are the key feature of the model, as they capture unobserved subject heterogeneity in the effect of experimental treatments on harvest. As indicated in (1) we model them to follow a multivariate normal distribution with mean  $\boldsymbol{\gamma}$  and variance matrix  $\boldsymbol{\Sigma}$ .

As discussed in more detail in Moeltner et al. (2008) for the dual reasons of truncation and parameter heterogeneity the mean-variance-equality constraint of the standard Poisson density is relaxed in the HDTP. Furthermore, the random coefficients in  $\lambda_{ip}$  introduce correlation between decisions made by a given individual over the  $P$  periods of the game. The likelihood function for the HDTP is given as

$$p(y_{ip} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, E_{\min} \leq y_{ip} \leq E_{\max}) = \prod_{i=1}^n \int_{\boldsymbol{\gamma}_i} \left( \prod_{p=1}^P \frac{\lambda_{ip}^{y_{ip}}}{y_{ip}! \left( \sum_{k=L}^U \frac{\lambda_{ip}^k}{k!} \right)} f(\boldsymbol{\gamma}_i | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) \right) d\boldsymbol{\gamma}_i \quad (2)$$

This would complete our model description in a classical framework. The analyst would choose a set of starting values for all model parameters and a computational algorithm to find the set of estimates that maximizes the likelihood function.

The Bayesian approach requires the specification of prior densities for all parameters. Its objective is to derive the joint posterior density, which, using Bayes' Rule, can be written as

$$p(\boldsymbol{\theta} | \mathbf{y}) = \frac{p(\mathbf{y} | \boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{y})} \propto p(\mathbf{y} | \boldsymbol{\theta}) p(\boldsymbol{\theta}) \quad (3)$$

where  $p(\mathbf{y} | \boldsymbol{\theta})$  is the sample likelihood,  $p(\boldsymbol{\theta})$  is the prior density, and  $p(\mathbf{y})$  is the marginal likelihood.

The latter construct will play an important role in model comparison, as discussed below, but can be treated as a normalization constant in the derivation of the posterior as it does not include  $\boldsymbol{\theta}$ . In other

words, to derive the posterior distribution we can focus on the product of the likelihood and the prior, which is “proportional” to the full posterior density as indicated in (3).<sup>1</sup>

For our HDTP framework our set of parameters, generically denoted as  $\theta$  above, includes  $\beta, \gamma$ , and the unique elements in  $\Sigma$ . Based on natural parameter restrictions and computational convenience we specify multivariate normal priors for  $\beta$  and  $\gamma$  and an inverse gamma prior for the diagonal elements of  $\Sigma$ , i.e.

$$\beta \sim mvn(\mu_\beta, V_\beta), \quad \gamma \sim mvn(\mu_\gamma, V_\gamma), \quad \Sigma_{jj} \sim ig(v_0, \varphi_0), \quad j = 1 \dots k_r \quad (4)$$

This implies that we are setting the covariances in  $\Sigma$  to zero. This restriction is generally not required for the HDTP. It is imposed at this stage solely to align our theoretical model exposition with our empirical application, where covariances between random treatment effect were found to be poorly identified, with posterior means close to zero. For a full (unrestricted)  $\Sigma$  a suitable and convenient prior would be the Inverse-Wishart (IW) density.

The core component of our Bayesian estimation is a Gibbs Sampler, i.e. a posterior simulator that draws sequentially and repeatedly from the following conditional probabilities:

$$\begin{aligned} p(\beta | y, X, H, \Sigma, \gamma_i), i = 1 \dots n, \quad p(\gamma | y, X, H, \Sigma, \gamma_i), i = 1 \dots n, \\ p(\Sigma_{jj} | \gamma, \gamma_i), i = 1 \dots n, j = 1 \dots k_r, \quad \text{and} \quad p(\gamma_i | y_i, X_i, H_i, \beta, \Sigma, \gamma), i = 1 \dots n \end{aligned} \quad (5)$$

Matrices  $X$  and  $H$  comprise, respectively, vectors  $x_i$  and  $h_{ip}$ , for all individuals and periods, and matrices  $X_i$  and  $H_i$  perform the same function for a given individual across all periods. Draws of  $\beta$  and  $\gamma_i$  require a Metropolis-Hastings (*MH*) sub-routine, as described in Chib et al. (1998) for the un-truncated hierarchical Poisson model. After a sufficient number of repetitions the conditional draws of  $\beta, \gamma$  and  $\Sigma$  will converge to the joint posterior distribution  $p(\beta, \gamma, \Sigma | y, X, H)$ . 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. We refer the reader to Moeltner et al.'s (2008) Appendix C for more details on these conditional densities and the posterior algorithm.

### *Marginal Likelihood and Model Selection*

As described in more detail in the next section we will estimate two different models, distinguished by the contents and hierarchical properties of treatment vector  $\mathbf{h}_{ip}$  in (1). These models are non-nested, i.e. one cannot be derived from the other by imposing a set of parameter restrictions. This would render a rigorous model comparison difficult in a classical estimation framework, but does not pose a problem in a Bayesian approach.

In Bayesian analysis, a “model” is defined by the prior densities chosen for its parameters and the form of its likelihood function. The basic building block for Bayesian model comparison is the marginal likelihood (*mLH*) for a given model  $m$ , generically expressed as

$$p(y|m) = \int_{\theta} p(y|\theta, m) p(\theta|m) d(\theta) \quad (6)$$

As discussed e.g. in Lancaster (2004, Ch. 2) the marginal likelihood or “prior predictive distribution” can be interpreted as describing what new data are expected to look like before they are collected, given the researcher’s choice of likelihood function and prior. It is then evaluated for collected data  $\mathbf{y}$ , conceptually by replacing the generic random variable  $y$  with data vector  $\mathbf{y}$  in (6).

The *mLH* can then be used to derive the ratio of model probabilities or “posterior odds ratio” for two competing specifications (e.g. Koop, 2004, Ch.1), given as

$$\frac{p(m_1|\mathbf{y})}{p(m_2|\mathbf{y})} = \frac{p(\mathbf{y}|m_1)p(m_1)}{p(\mathbf{y}|m_2)p(m_2)} \quad (7)$$

where  $p(\mathbf{y}|m)$  is the model- conditioned marginal likelihood evaluated at the observed data, and  $p(m)$  indicates the prior model probability, which is assigned by the researcher at the onset of the analysis. As is evident from (7) 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.

The computational difficulty in deriving *BF*'s lies in the derivation of the *mLH*, which in most cases does not have a closed-form solution. Instead, it has to be evaluated via approximation or simulation.<sup>2</sup> In this application we use the framework outlined in Chib (1995) and Chib and Jeliazkov (2001) to derive the value of the *mLH* via simulation. This method starts by re-writing Bayes' Rule in (3) as

$$p(\mathbf{y} | m) = \frac{p(\mathbf{y} | \boldsymbol{\theta}, m) p(\boldsymbol{\theta} | m)}{p(\boldsymbol{\theta} | \mathbf{y}, m)}, \quad (8)$$

Since the left hand side of (8) does not contain  $\boldsymbol{\theta}$ , the relationship must hold for any choice of parameter values. This property is deemed "marginal likelihood identity" in Chib (1995). In practice,  $p(\mathbf{y} | m)$  is usually evaluated at a point of high posterior density for  $\boldsymbol{\theta}$  such as the posterior mean. We evaluate the *mLH* in log form at the posterior mean for each model, i.e. we use

$$\log(p(\mathbf{y} | m)) = \log(p(\mathbf{y} | \bar{\boldsymbol{\theta}}, m)) + (\log(p(\bar{\boldsymbol{\theta}}, m)) - \log(p(\bar{\boldsymbol{\theta}} | \mathbf{y}, m))) \quad (9)$$

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 (9) 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. While the first two densities on the right hand side of (9) are generally known, the evaluation of the full posterior distribution at  $\bar{\boldsymbol{\theta}}$  requires additional reduced Gibbs Samplers, as described in Chib (1995) and Chib and Jeliazkov (2001).

### Posterior Predictions

The Gibbs sampler described above generates  $r=1 \dots R$  draws of parameters. These draws can then be used to derive posterior predictive distributions (*PPDs*) for outcomes of interest for any combination of individual characteristics and experimental treatments. Given their importance in informing resource management policies we will focus on expected harvest levels in our application. We denote a specific predictive setting for our regressors as  $\mathbf{x}_f$  and  $\mathbf{h}_f$ , respectively. Conceptually, the *PPD* for expected harvest,  $E(y | \mathbf{x}_f, \mathbf{h}_f)$ , can then be expressed as

$$p(E(y | \mathbf{x}_f, \mathbf{h}_f)) = \int_{\boldsymbol{\theta}} \left\{ \int_{\gamma_i} \left( \sum_{k=E_{\min}}^{E_{\max}} \frac{\lambda_{if}^k}{k!} \right)^{-1} \sum_{y=E_{\min}}^{E_{\max}} \frac{y \lambda_{if}^y}{y!} \right\} f(\gamma_i | \boldsymbol{\gamma}, \boldsymbol{\Sigma}) d\gamma_i \Bigg\} p(\boldsymbol{\theta} | \mathbf{y}, \mathbf{X}, \mathbf{H}) d\boldsymbol{\theta} \quad (10)$$

where  $\lambda_{if} = \mathbf{x}'_f \boldsymbol{\beta} + \mathbf{h}'_f \boldsymbol{\gamma}_i$ , and  $\boldsymbol{\theta}$  comprises all model parameters. For the HDTP the expectation over harvest is to be understood as the expectation over the truncated Poisson kernel, conditional on random effects  $\gamma_i$  and a given draw of model parameters  $\boldsymbol{\theta}$ . The posterior predictive simulator then removes the latter conditionalities by estimating expected harvest for repeated draws of  $\gamma_i$  and  $\boldsymbol{\theta}$ . The details of this approach are given in Moeltner et al. (2008, Appendix D).

In our application we compute separate *PPD*'s for expected harvest for two geographically disparate villages. A *PPD* for *combined* harvest across villages can then be derived in straightforward fashion via the convolution of the village-specific predictive distributions. Denoting harvest levels and predictive densities for the two areas as  $y_1$  and  $y_2$ , and  $p_1(\cdot)$  and  $p_2(\cdot)$ , respectively, the *PPD* for combined expected harvest can be formally expressed as

$$p(E(y_1 + y_2 | \mathbf{x}_{f,1}, \mathbf{h}_{f,1}, \mathbf{x}_{f,2}, \mathbf{h}_{f,2})) = \int_{E(y_1 | \mathbf{x}_{f,1}, \mathbf{h}_{f,1})} p_1(E(y_1 | \mathbf{x}_{f,1}, \mathbf{h}_{f,1})) p_2(S - E(y_1 | \mathbf{x}_{f,1}, \mathbf{h}_{f,1})) dE(y_1 | \mathbf{x}_{f,1}, \mathbf{h}_{f,1}) \quad (11)$$

where  $\mathbf{x}_{f,j}$  and  $\mathbf{h}_{f,j}$ ,  $j=1,2$  are, respectively, the predictive settings for demographic characteristics and institutional treatments at either location, and  $S = E(y_1 + y_2 | \mathbf{x}_{f,1}, \mathbf{h}_{f,1}, \mathbf{x}_{f,2}, \mathbf{h}_{f,2})$ . In practice draws from

the convoluted *PPD* can be obtained by adding draws from  $p_1(E(y_1 | \mathbf{x}_{f,1}, \mathbf{h}_{f,1}))$  to draws from  $p_2(E(y_2 | \mathbf{x}_{f,2}, \mathbf{h}_{f,2}))$  via a “complete combinatorial” or via random sampling, as discussed e.g. in Poe et al. (2005). It should be noted that this combination of location-specific *PPD*’s provides considerable flexibility in exploring the total effect on resource extraction of *different mixes of policy tools* at different locations. We will illustrate this strategy in our empirical section.

#### IV) Empirical Application

##### *Data*

To illustrate our estimation framework we use data from CPR field experiments conducted in 2004 in several artisanal fishing communities of Colombia. These experiments are described in detail in Velez et al. (Forthcoming(a)) and Velez et al. (Forthcoming(b)). For this application we focus on two of these communities, Ensenada de Tumaco on the Pacific coast (abbreviated as PAC in the following), and La Dorada in the Magdalena River’s watershed in the country’s interior (henceforth abbreviated as MAG). We chose these two areas because of their different histories with respect to the institutional regulation of local fisheries. Specifically, the shrimp fishery of Ensenada de Tumaco has a long-standing tradition of government regulations as well as informal regulations. In contrast, government regulation is relatively weak in the Magdalena watershed, so local fishermen have largely relied on informal associations to manage their fresh water fishery. These pronounced differences in institutional experience lead to interesting differences in experimental outcomes, as we will discuss below. Some basic sample statistics for these two communities are given in Table 1.

In each area subjects played several rounds of a CPR game in groups of five players. In each round the participants had to decide to harvest between 1 and 9 integer 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 payoff table and its underlying profit function are given in Velez et al. (Forthcoming(a, b)) .

The game was administered under different institutional “treatments”. For this application we focus on three of these treatments: “Treatment 1” ( $T_1$ ), a “Limited Access” version of the game without government intervention or communication (although with an upper bound of nine units of harvest), “Treatment 2” ( $T_2$ ), a harvest quota of 2 units, with a 10% audit probability and a low penalty, and “Treatment 3” ( $T_3$ ), the opportunity to openly communicate with the other group members prior to each round of decision-making.<sup>3</sup> Further details on the implementation of the experiment are given in Velez et al. (Forthcoming (a,b)). All 20 individuals (four groups) in each of the two communities included in our sample played ten rounds of the Limited Access game. For the second set of ten rounds, ten individuals (two groups) in each village played ten rounds under  $T_2$ , with the other half of participants playing an equal number of rounds under  $T_3$ . Thus, our total sample includes 800 observations from 40 individuals on harvest decisions, 400 for  $T_1$  and 200 for each of  $T_2$  and  $T_3$ .

### *Model specification*

As mentioned above we estimate and compare two models with different conditional mean functions for each fishing community. For both models the vector of individual characteristics  $\mathbf{x}_i$  includes “gender” (1 = female) and “years of education”<sup>4</sup>. Model 1 ( $M_1$ ) has no constant term but includes 0/1 indicators for each of the three treatments. It also assigns random coefficients to each treatment, and thus allows for *institution-specific* unobserved heterogeneity in observed harvest decisions.

The conditional mean function for this model is thus given as

$$\log \lambda_{ip} = \mathbf{x}_i' \boldsymbol{\beta} + \gamma_{i1} T_1 + \gamma_{i2} T_2 + \gamma_{i3} T_3 \quad \text{with}$$

$$\begin{bmatrix} \gamma_{i1} \\ \gamma_{i2} \\ \gamma_{i3} \end{bmatrix} \sim n \left( \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix} \right) \quad (12)$$

This implies that the variance of the logged mean function varies with treatments, i.e.

$$V(\log \lambda_{ip} | T_j) = \sigma_{jj}, j = 1 \dots 3 \quad (13)$$

Accordingly, the variance for the truncated, un-logged expectation will also be treatment-specific, as will be evident in our posterior predictive distributions below.

Model 2 ( $M_2$ ) is similar to the basic random effects specification that has been frequently used in other applications of multi-period games to capture unobserved heterogeneity in individual decisions (e.g. Carpenter, 2004, Moffat, 2005, Velez et al., Forthcoming (a), Velez et al., Forthcoming (b)). It includes a random constant term and fixed effects for treatments two and three. Treatment one becomes the underlying baseline category. Formally:

$$\begin{aligned} \log \lambda_{ip} &= \mathbf{x}_i' \boldsymbol{\beta} + \gamma_2 T_2 + \gamma_3 T_3 + \mu_i \quad \text{with} \\ \mu_i &\sim n(0, \sigma_\mu^2) \end{aligned} \quad (14)$$

The important difference to  $M_1$  is that in this case unobserved heterogeneity is *not systematically linked to treatments* (or any other explanatory variable). The two models are not nested since the derivation of  $M_2$  from  $M_1$  requires the simultaneous restriction of  $\sigma_{11} = \sigma_{22} = \sigma_{33} = 0$  and the addition of the random constant term with variance  $\sigma^2$ . Note that this variance is based on all 20 observations associated with individual  $i$ , while the treatment-specific variances in  $M_1$  are each based on the ten observations associated with a given treatment. Model 2 also implies that the variance of the logged mean function is equal under all treatments, i.e.

$$V(\log \lambda_{ip} | T_j) = \sigma_\mu^2, j = 1 \dots 3, \quad (15)$$

which translates into an equal variance for the truncated un-logged expectation over treatments.

### *Prior refinement*

We first estimate both models using the full sample and the following vague but proper priors:  $\boldsymbol{\mu}_\beta = 0, \mathbf{V}_\beta = 100, \boldsymbol{\mu}_\gamma = 0, \mathbf{V}_\gamma = 100, \eta_0 = \varphi_0 = 1/2$ . Both 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, 1992)'s convergence diagnostic (CD). We label this approach the “*Full Sample*” strategy.

Our Bayesian framework offers an interesting alternative option in utilizing our data. Specifically, we can use some of our data to generate refined priors which can then be combined with the remaining data for the main analysis, along the lines described in Moeltner et al. (2007). In our application, for example, we split the sample into the first five rounds for a given individual and treatment, and the second set of five rounds. We use the first five rounds in conjunction with the diffuse priors given above to generate posterior distributions for all parameters. We then employ the moments of these empirical distributions to replace the initial set of vague prior moments. We then combine these refined priors with the remaining data to generate the final set of posterior densities. The first step yields variance settings for  $\mathbf{V}_\beta$  and  $\mathbf{V}_\gamma$  in the 0.2 to 0.8 range, and settings for the inverse-gamma shape and scale parameters between 1.5 and 2, and 0.1 to 0.4, respectively. We deem this approach the “*Split-Sample*” strategy.

This Split-Sample strategy might be attractive in settings where the researcher wants to assign somewhat lower weight to the first few rounds of a game, (for example due to hypothesized presence of pronounced learning effects or other undesirable noise), but is reluctant to discard early-round results outright. For our data, learning or other temporal effects have been found to be of secondary importance in previous analyses (Moeltner et al., 2008, Velez et al., Forthcoming (a), Velez et al., Forthcoming(b)). However, the Split-Sample approach can still provide overall efficiency gains over the Full Sample strategy if the gain in prior precision outbalances the loss in actual data in the second stage. This is clearly apparent in our estimation results, as documented below.

### *Estimation results*

Table 2 captures the components of the log-*mLH* expression in (9) for both models and regions. The last column provides the logged Bayes Factor, which refers to the difference of the log *mLH* of the

more likely model to the  $\log mLH$  of the other specification. The upper block of the table depicts results for the Full Sample approach, while the lower block captures results for the Split-Sample strategy. As is evident from the Table, the unconstrained model ( $M_1$ ) receives substantially higher relative posterior probability than the restricted version ( $M_2$ ) in all cases. Using the interpretation thresholds for  $BFs$  suggested in Kass and Raftery (1995) there is “decisive” ( $\log BF > 11.5$ ) evidence that  $M_1$  is more likely to have generated the observed data than the random constant-only model ( $M_2$ ) in all cases. As can be seen from inspection of the individual components that feed into the  $\log mLH$  value (columns three to five) the observed differences in  $mLH$  are largely driven by pronounced differences in sample likelihood values. We thus conclude that regardless of estimation approach (i.e. Full Sample vs. Split Sample) the likelihood function for the unconstrained model with treatment-specific heterogeneity fits the data by orders of magnitude better than the likelihood function for the random-constant model. However, to further highlight the implications of ignoring institution-specific heterogeneity we will continue to report results for both models throughout the remainder of this section.

Table 3 shows the posterior mean and standard deviation for all parameters included in both models. The results are again reported by region (PAC = first two columns, MAG = last two columns) and estimation strategy (left table = Full Sample, right table = Split Sample). The first block of rows shows results for the fixed effects, i.e. the elements of  $\beta$  for  $M_1$  and  $\beta$  and  $\gamma$  for  $M_2$ , the second block corresponds to the means of random effects, i.e. the elements of  $\gamma$  for  $M_1$  and the constant term for  $M_2$ , and the third block gives results for the random effect variances, i.e. the diagonal elements of  $\Sigma$  for  $M_1$  and the variance of the random constant,  $\sigma_\mu^2$ , for  $M_2$ .

For both models the estimated posterior means for the fixed effects “female” and “education” can be interpreted as the expected marginal effects of these regressors on the logged (un-truncated) expected extraction. The posterior means for treatment effects can be loosely interpreted as the log of (un-truncated) expected harvest for a male participant with no schooling. They are captured in the “random effect means”- block of the table for  $M_1$ . This block also contains the mean of the random constant for

$M_2$ , which corresponds to the posterior mean effect of  $T_1$ , the baseline category in that model. In contrast, the posterior means for the effects of  $T_2$  and  $T_3$  are shown in the “fixed effects” block for the restricted model. Since  $M_2$  includes a constant term, these posterior means are to be interpreted as the *differential* effect for each treatment compared to the Limited Access scenario. For all parameters the posterior standard deviations, provided in parentheses beneath each posterior mean, indicate the level of precision with which these effects are estimated.

There are three important results that flow from Table 3: (i) both models agree on the ranking of treatment effects within each region, (ii) the random effect variances for all three treatments in Model 1 have high posterior precision, and their posterior means clearly differ across treatments, and (iii) the Split Sample models exhibit higher posterior efficiency for most parameters.

With respect to finding (i) we note that the posterior means of expected harvest levels are highest for the Limited Access scenario ( $T_1$ ) for all regions and models. For PAC the institutional treatment  $T_2$  has a decidedly stronger effect on harvest reduction than communication. This picture is reversed for MAG, where  $T_3$  induces the lowest mean extraction levels followed by the quota-plus- penalty regulatory scenario. As discussed in more detail in Moeltner et al. (2008) these results are intuitively sound considering the historical institutions for harvest regulation in each community. The new insight in this study is that the unrestricted and restricted models portray the same ranking across treatments. Thus, if such a ranking was our sole research objective, the proper modeling of treatment-specific heterogeneity would be of secondary importance for our application.

However, important additional insights about the effects of each institutional design can be gained from the unrestricted model. As summarized in finding (ii) above and depicted in the bottom block of the table for  $M_1$ , the three treatments produce clearly distinct posterior means for random effect variances. First, our results clearly indicate that allowing for communication ( $T_3$ ) reduces random effect variances compared to external regulation ( $T_2$ ). In other words, communication appears to be better suited to reduce heterogeneous noise in decision making and foster the coordination of individual strategies than the weakly enforced quota regime in  $T_2$ . This result holds for both regions and estimation strategies.

Second, the Limited Access treatment produces the lowest level of heterogeneous noise for both regions under the Full Sample approach. However, as discussed in Moeltner et al. (2008), this result may be an artifact related to the experimental design, which produces twice as many observations for decisions under  $T_1$  than for outcomes under  $T_2$  and  $T_3$ . This allows for the diffuse priors used in the Full Sample model to have a relatively stronger impact on the posterior results for the variances associated with the latter two treatments. Not surprisingly, this sample size effect is diminished under the Split Sample approach, which feeds substantially tighter priors for random effect variances into the posterior simulator. As is evident from the first column of the Split-Sample table, this actually lowers heterogeneity noise under  $T_3$  below the  $T_1$ -level. Since many SD experiments produce unbalanced sets of observations for different policy treatments, the Split-Sample's ability to diminish the asymmetric influence of vague priors on posterior results may be another attractive reason to employ this framework.

Building on finding (iii) above we also note that the Split Sample approach generates very similar posterior means compared to the Full Sample results, but exhibits lower posterior standard deviations for most parameters, regardless of model or region. Thus, the tradeoff between reduced sample size and increased prior precision implicit in the Split Sample framework resulted in an overall efficiency gain for the final set of results for our application. Naturally, it is possible that further efficiency gains could be achieved by changing our split-up over rounds from 5/5 to, say, 4/6, 6/4, or other proportions. It would be interesting to explore this question of "optimal sample split" in future research.

Next, we compare Models 1 and 2 based on their posterior predictive distributions (*PPDs*) for expected harvest, as described in the previous Section. We derive 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 "gender" indicator, using the observed proportion of females as weights. 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.<sup>5</sup>

Table 4 shows the posterior means and standard deviations for predicted expected harvest for all models and regions. As in Table 2 the first half of the table reports results corresponding to the Full Sample approach, while the bottom half shows results for the Split-Sample strategy. The first row of each sub-table depicts corresponding sample statistics, evaluated over all ten periods for the Full Sample, and over the last five periods of a given treatment for the Split Sample. In general, the restricted and unrestricted models produce posterior means of comparable magnitude to one another and to the sample statistics. As expected (and discussed in more detail in Moeltner et al. , 2008),  $T_2$  induces the lowest expected harvest for the Pacific community, while  $T_3$  yields the lowest extraction level for Magdalena.

From this study's perspective the key result captured in Table 4 is the pronounced difference in posterior standard deviations between the constrained and unconstrained models for treatments two and three. Specifically, the restricted model  $M_2$  produces substantially smaller standard deviations for these treatments than the unrestricted specification. While this might be naïvely interpreted as “superior efficiency”, it is a direct result of ignoring individual heterogeneity in subjects' responses to these institutional treatments. We will discuss this effect further in our graphical analysis below. True efficiency gains, however, are again delivered by the Split-Sample approach compared to the Full Sample strategy, as the former generates slightly lower posterior standard deviations than the latter for most of the PPD's captured in the table. This gain in predictive efficiency is, of course, directly related to the Split Sample's higher precision in the estimation of the underlying parameters as discussed above.

Figure 1 depicts the *pdf* and *cdf* for all simulated *PPDs* of expected harvest. All sub-figures are based on the Split-Sample results.<sup>6</sup> While the distribution under  $T_1$  is similar for models  $M_1$  and  $M_2$ , it is apparent from the Figure that Model 2 produces misleadingly tight and symmetric distributions for  $T_2$  in the Pacific application, and for  $T_2$  and  $T_3$  for the Magdalena sample. Model 1 reveals that the true *PPDs* under these treatments are skewed to the right, with considerable probability mass over the lowest harvest tier. Furthermore, it is evident only from  $M_1$  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 completely absent for the restricted model.

Figure 2 shows the combined expected harvest under the four possible mixes of treatments two and three across regions. As discussed in the previous section these combined *PPDs* are generated by convoluting the corresponding region-specific predictive densities.<sup>7</sup> As for Figure 1, all sub-figures are based on Split-Sample estimation. For comparison purpose, each sub-figure also shows the *PPD* for expected extraction under the Open-Access scenario (solid line). The policy relevance of these figures lies in their ability to depict optimal harvest-reducing strategies under a single umbrella policy (same treatment at both regions), and under region-specific policies. This is important as in a given real-world context legislative or budgetary constraints may impose one or the other approach. The figure clearly indicates that the ideal scenario under unrestricted policy choice would be to implement  $T_2$  for PAC and  $T_3$  for MAG (lower right sub-figure). This is not surprising given our previously discussed results on parameter estimates and region-specific *PPDs*.

The figure also offers a clear prescription if a single policy instrument has to be chosen in both regions. Comparing the top two sub-figures, it is clear that both  $T_2$  and  $T_3$  induce the same upper bound of approximately 14 units, but  $T_2$  allocates substantially more mass to the lowest extraction tier. This implies that the relatively stronger effect of  $T_2$  to reduce harvest in the PAC region outweighs the comparative advantage of  $T_3$  in MAG when the objective is to reduce combined extraction.

#### IV) Conclusion

In this chapter we illustrate the importance of allowing for treatment-specific subject heterogeneity in the analysis of data from Social Dilemma Games. Using a Bayesian estimation framework and data from common-pool resource field experiments in rural Colombia, we find that a naïve model that captures unobserved heterogeneity only in generic terms produces severely misleading results with respect to the variability and distribution of predicted outcomes under different treatments.

We also demonstrate how early-round data can be used to derive refined priors that, when combined with the remaining data, can enhance the posterior efficiency of final results. We further show how experiment-specific predictive results can be combined via the method of convolution to generate densities of broader predictive constructs. This can be an important tool to inform region-wide resource policy under different administrative constraints of treatment allocation over sub-regions.

In addition to highlighting methodological issues and tools related to treatment-specific subject heterogeneity, this paper also advances the understanding of potential differences in effectiveness between externally-imposed regulations and informal rules. In particular, it lends support to the intuitive notion that communication serves as an effective device for coordinating individual decisions, thereby reducing heterogeneous noise. On the other hand, the bi-modal outcome of  $T_2$  in Magdalena (M1 in Figure 1) shows that although external regulations may reduce harvests on average, this may be the result of vastly different individual responses to that treatment. Although average responses to the two treatments may be similar, the paths through which this is accomplished can be fundamentally different. Model 1 draws this point out, whereas this conclusion would be missed with Model 2.

Naturally, all of our results are based on quantity-neutral “extraction choices” from stylized games. Since the preference structure underlying the pay-off table was exogenously imposed by the researcher, it is difficult to link the resulting frame-constrained choices to actual expected harvest levels in units of physical quantities. To make meaningful predictions regarding actual resource use based on experimental results, the experimental design needs to reflect real-world preference structures and production processes. We believe this would be a natural next step to better align field experiments with actual policy prescriptions. It is our hope that with this chapter we have contributed to paving the path towards this important future extension of field experiments with focus on natural resource management.

Notes

<sup>1</sup> The reader is referred to Koop (2004), Lancaster (2004) or Koop et al. (2007) for further details on the fundamental concepts and method in Bayesian estimation.

<sup>2</sup> For a comparative review of alternative approaches to approximate the marginal likelihood see Han and Carlin (2001).

<sup>3</sup> From a game-theoretic perspective the expected penalties under regulatory treatment  $T_2$  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>4</sup> We also included subjects' age in preliminary specification but found no measurable effects for this regressor.

<sup>5</sup> Our predictive draws are generated as follows: We first thin the original sequence of parameter draws flowing from the Gibbs Sampler by a factor of ten to reduce simulation-induced autocorrelation. For each of the resulting 1200 draws of parameters, we draw 100 values for random effects. This yields a total of 120,000 predictive draws for expected harvest.

<sup>6</sup> The Full Sample produces very similar graphs. They are available from the authors upon request.

<sup>7</sup> To conserve on computational memory, we generate these convolutions by recouping and convoluting the 1200 predictions for expected harvest corresponding to the thinned parameter sequence from the original Gibbs Sampler plus a single draw of random effects.

References:

- 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. 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.
- Beckenkamp, M. and A. Ostmann. 1999. Missing the target? Sanctioning as an ambiguous structural solution. In *Resolving social dilemmas*, M. Foddy, M. Smithson, S. Scheneider and M. Hogg (eds.). Psychology Press: Philadelphia.
- Bischoff, I. 2007. Institutional choice versus communication in social dilemmas - an experimental approach. *Journal of Economic Behavior and Organization* **62**: 20-36.
- Cardenas, J. C., J. Stranlund and C. Willis. 2000. Local environmental control and crowding-out. *World Development* **28**: 1719-1733.
- Cardenas, J. C. and E. Ostrom. 2004. What do people bring into the game? Experiments in the field about cooperation in the commons. *Agricultural Systems* **82**: 307-326.
- 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., 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.
- Fischbacher, U., S. Gächter and E. Fehr. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* **71**: 397-404.
- 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.

- 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.
- 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:
- List, J. A. 2006. Field experiments: A bridge between lab and naturally occurring data. *B.E. Journal of Economic Analysis & Policy* **6**: 1-45.
- Moeltner, K., K. Boyle and R. Paterson. 2007. Meta-analysis and benefit transfer for resource valuation - addressing classical challenges with Bayesian modeling. *Journal of Environmental Economics and Management* **53**: 250 - 269.
- Moeltner, K., J. J. Murphy, J. K. Stranlund and M. A. Velez (2008). Processing data from social dilemma experiments: A Bayesian comparison of parametric estimators, University of Nevada, Reno, Joint Economics Working Paper Series # 07013.
- Moffat, P. G. 2005. Stochastic choice and the allocation of cognitive effort. *Experimental Economics* **8**.
- Ostmann, A. 1998. External control may destroy the commons. *Rationality and Society* **10**: 103-122.
- Ostrom, E, R. Gardner and J. Walker. 1994. *Rules, games, and common-pool resources*. University of Michigan Press: Ann Arbor
- 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.
- Poe, G. L., K. L. Giraud and J. B. Loomis. 2005. Computational methods for measuring the difference of empirical distributions. *American Journal of Agricultural Economics* **87**: 353-365.
- 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.

Table 1: Sample Statistics

	Pacific		Magdalena	
	mean / percent	std	mean	std
age	43.2	11.6	42.7	15.4
years of formal education	4.1	2.3	4.3	2.9
female	10.0%	-	20.0%	
lived in community ten years or longer	95.0%	-	93.0%	-
fishing is the main household activity	94.0%	-	87.0%	-
extraction , periods 1-5	5.1	2.4	5.0	2.5
extraction, periods 6-10	4.9	2.6	4.9	2.7

Table 2: Model Comparison

Full Sample						
	# param.	logLH	log(prior)	log(post.)	log(mLH)	logBF*
Pacific						
Model						
M1	8.00	-1471.19	-20.57	8.50	-1500.27	173.12
M2	6.00	-1650.51	-18.43	4.45	-1673.39	-
Magdalena						
Model						
M1	8.00	-1509.88	-20.34	6.96	-1537.19	226.17
M2	6.00	-1741.25	-17.43	4.67	-1763.36	-
Split Sample						
	# param.	logLH	log(prior)	log(post.)	log(mLH)	logBF*
Pacific						
Model						
M1	8.00	-682.556	6.590	10.657	-686.623	152.05
M2	6.00	-838.526	5.677	5.822	-838.672	-
Magdalena						
Model						
M1	8.00	-673.805	3.343	7.854	-678.316	195.85
M2	6.00	-874.083	5.261	5.347	-874.169	-

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

Table 3: Estimation Results for Most Likely Sub-models, all Regions

	Full Sample				Split Sample			
	Pacific		Magdalena		Pacific		Magdalena	
	M1	M2	M1	M2	M1	M2	M1	M2
	mean	mean	mean	mean	mean	mean	mean	mean
	(std)	(std)	(std)	(std)	(std)	(std)	(std)	(std)
<u>Fixed Effects</u>								
female	0.118	0.048	0.226	0.237	0.078	0.048	0.244	0.250
	(0.162)	(0.184)	(0.144)	(0.169)	(0.125)	(0.134)	(0.115)	(0.125)
education	-0.230	-0.234	-0.027	0.082	-0.279	-0.272	0.005	0.049
	(0.221)	(0.275)	(0.173)	(0.200)	(0.148)	(0.166)	(0.123)	(0.136)
T2	-	-0.904	-	-0.490	-	-0.892	-	-0.481
	-	(0.057)	-	(0.054)	-	(0.057)	-	(0.054)
T3	-	-0.242	-	-0.537	-	-0.260	-	-0.563
	-	(0.049)	-	(0.055)	-	(0.050)	-	(0.055)
<u>RE Means</u>								
const	-	2.020	-	1.761	-	2.009	-	1.813
	-	(0.131)	-	(0.115)	-	(0.075)	-	(0.078)
T1	2.017	-	1.802	-	2.016	-	1.802	-
	(0.113)	-	(0.103)	-	(0.077)	-	(0.074)	-
T2	1.010	-	1.372	-	1.037	-	1.344	-
	(0.201)	-	(0.255)	-	(0.140)	-	(0.170)	-
T3	1.697	-	1.143	-	1.664	-	1.126	-
	(0.135)	-	(0.169)	-	(0.089)	-	(0.120)	-
<u>RE Variances</u>								
const	-	0.105	-	0.159	-	0.079	-	0.136
	-	(0.029)	-	(0.042)	-	(0.024)	-	(0.039)
T1	0.116	-	0.118	-	0.105	-	0.129	-
	(0.035)	-	(0.034)	-	(0.041)	-	(0.047)	-
T2	0.587	-	1.072	-	0.408	-	1.023	-
	(0.241)	-	(0.436)	-	(0.180)	-	(0.432)	-
T3	0.146	-	0.310	-	0.068	-	0.302	-
	(0.057)	-	(0.125)	-	(0.033)	-	(0.134)	-

RE = random effects

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

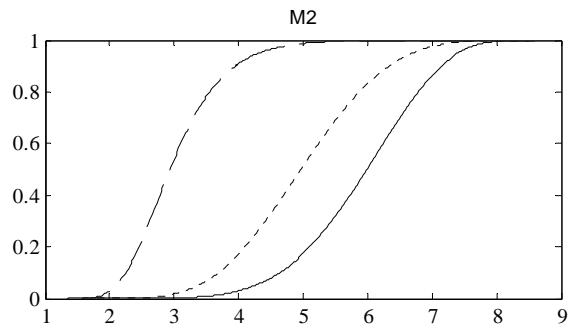
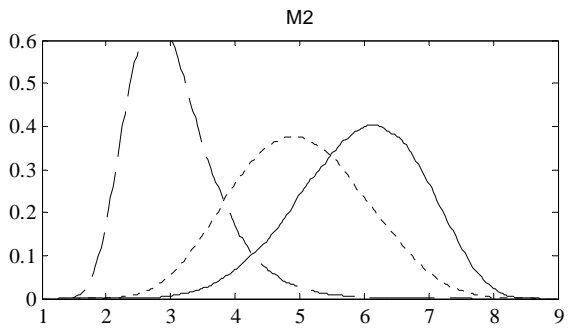
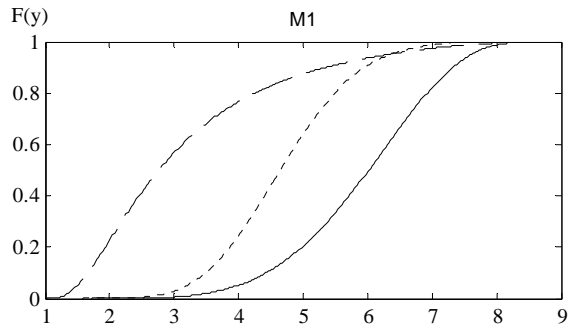
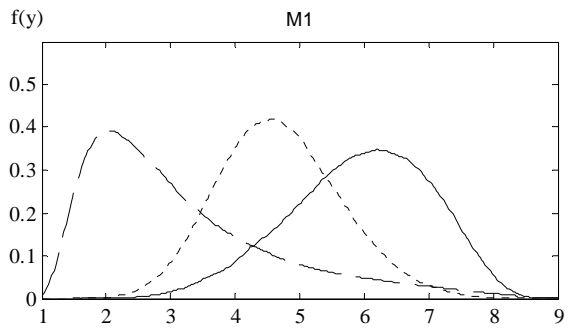
Table 4: Observed and Predicted Extraction Levels

	Pacific			Magdalena		
	T1	T2	T3	T1	T2	T3
<b>Full Sample</b>						
Sample						
sample mean	5.920	3.260	4.875	5.620	4.355	3.575
Model Predictions						
M1						
post. mean	6.008	3.295	4.909	5.667	4.530	3.696
(post. std)	(1.143)	(1.744)	(1.310)	(1.180)	(2.379)	(1.512)
M2						
post. mean	5.986	3.102	5.153	5.685	4.088	3.935
(post. std)	(1.097)	(0.824)	(1.156)	(1.311)	(1.239)	(1.214)
<b>Split Sample</b>						
Sample (last 5 periods)						
sample mean	5.770	3.350	4.520	5.553	4.670	3.680
Model Predictions						
M1						
post. mean	5.935	3.161	4.689	5.710	4.475	3.682
(post. std)	(1.101)	(1.473)	(0.952)	(1.213)	(2.313)	(1.470)
M2						
post. mean	5.929	3.033	4.990	5.825	4.219	3.958
(post. std)	(0.957)	(0.686)	(1.001)	(1.226)	(1.191)	(1.149)

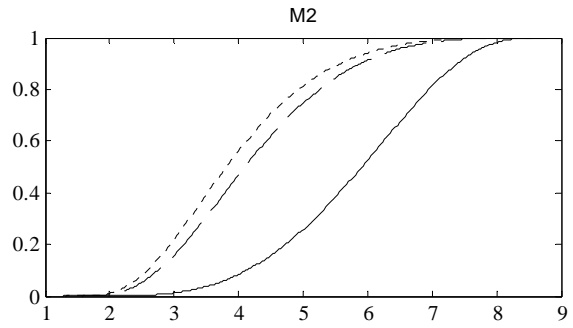
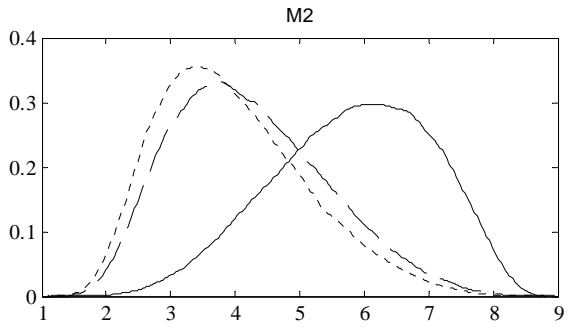
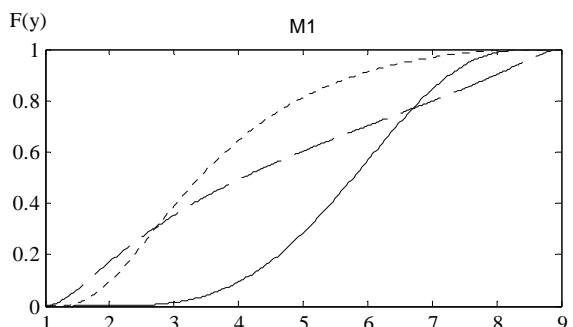
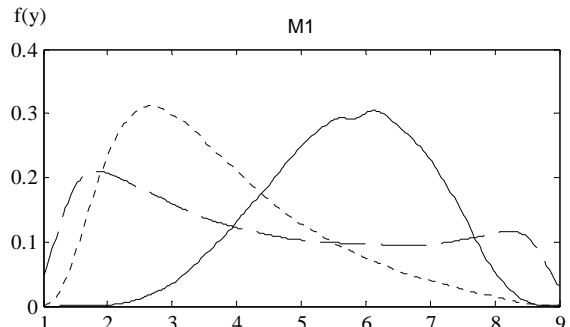
---

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

**Figure 1: Predictive pdf and cdf for Expected Harvest Pacific**



*Magdalena*



Legend: — T1    - - - T2    ..... T3

Figure 2: Predictive pdf for Combined Expected Harvest Under Different Treatment Combinations

