

Measuring Culpability

Katherine Barnes¹

James E. Rogers College of Law, University of Arizona

David Sloss

Santa Clara University School of Law

and

Stephen Thaman

Saint Louis University School of Law

preliminary and incomplete

do not cite or quote without permission

1 Introduction

Measuring the culpability of a crime (or criminal) has long been a goal of the criminal justice system. Legislators enact particular punishments (generally, a range of punishments) on the basis of the relative perceived culpability of different crimes; judges meet out actual sentences with the same type of proportionality in mind; prosecutors decide whom to charge, and for what crime based, in theory, on this issue. Indeed, all of these actors are supposed to be making these judgments on the basis of culpability for the crime: how bad the crime was, and, concurrently, how culpable the defendant was for the crime.

Thus, measuring culpability (judging how heinous a crime was) in order to calibrate the right amount of punishment happens, intuitively, every day in the criminal justice system.

¹Corresponding author. The authors would like to thank Marc Miller, David Baldus, Catherine Grosso, Jeffrey Fagan, Christopher Slobogin, Margo Schlanger, Carol Rose, Jack Chin, Doug Berman, and participants at colloquia at Duke Law School, Conference on Empirical Legal Studies, and University of Virginia for helpful comments. Barnes would also like to acknowledge the assistance of Professor John Hobbins at the McGill Law Library, for kindly providing work space and access to materials while this manuscript was being completed.

More deliberately, researchers have also sought to measure culpability, in order to control for different case characteristics when investigating how judges, prosecutors and juries make decisions about the heinousness of a crime. Theoretically, prosecutors, legislators, and juries are supposed to voice the conscience of the community, representing the overall public sentiment about each crime. But many have voiced concern that these decision-makers are too easily swayed by factors that are not appropriate, and should not be relied upon in making moral judgments about a person's actions. Chief among these factors is the race of the defendant, or of the victim(s). More recently, the geographic location of the crime (which is clearly related to race as well) has come under fire as an improper and arbitrary basis for decisionmaking.

In order to determine whether actors in the criminal justice system - generally, judges, juries, or prosecutors - are relying on race or other improper factors,² one has to control for the culpability of a crime. Without this, a researcher determine whether the mix of cases differs in important ways across race or geography.³ Indeed, researchers have generally attempted to control for culpability by controlling for individual case characteristics, under the theory that these case characteristics create the difference in culpability. So, for example, researchers investigating prosecutorial decision-making with regard to the death penalty have used case characteristics such as the number of victims, special victim characteristics, categories of crime, and additional crimes committed at the same time to control for culpability. Researchers have also included defendant characteristics, such as criminal history, to control for culpability. In each of these cases, researchers are indirectly measuring culpability by controlling for it (at least partially) in regressions. However, because these methods do not directly measure culpability, there is always the danger that some important case characteristic was not controlled for, making causal inference difficult using these methods. Finally, researchers have used a "salient factors" scale, adding points for

²I recognize that there may be significant disagreement about what constitutes a proper or an improper factor in judgment, based upon differing views of the proper role of punishment in the criminal justice system. However, this paper tackles race as a factor, about which there is more substantial agreement.

³There may be situations where the case mix across actors (judges, for example) is reasonably assumed to be random, and therefore the specific nature of each crime may be unnecessary to measure. But even in these situations, there is an argument that different sub-populations of defendants may committed different levels of crimes, on average, thus skewing any results that differentiate between groups.

aggravating circumstances and subtracting points for mitigating ones. Choosing how many points and whether positive or negative is, however, a somewhat subjective practice.

This article posits a new way to measure culpability, also using a survey of individuals to measure culpability, and two different methods to estimate the culpability measures that allows the values to vary continuously. Allowing culpability to vary continuously is advantageous because it more closely matches reality, and opens the door to multiple ways to control for culpability more completely in other models. After introducing the model and explaining the survey, we perform a case study on Missouri homicides, contrasting the use of prior methods with our methods to control for culpability. If, as one Missouri prosecutor suggested, the wide disparities in charging decisions and conviction rates are based upon the individual culpability of the case, then these disparities should disappear when one controls for culpability in the model. We end by comparing the new estimates to other rankings in the survey. In future work, we also plan to investigate the distribution of culpability across cases, correlate these to case characteristics. An additional byproduct of our model is the estimation of individual preference positions regarding culpability and punishment; we can therefore investigate the distribution of preferences across our sample as well.

This remaining portion of this Article is organized as follows: Section 2 introduces the IRT model, and our Bayesian completion of the model. Section ?? describes the data and data collection methods. Section ?? provides preliminary evidence that the model is well-identified and can differentiate between cases and individual preferences. Section ?? provides the case study of the Missouri death penalty, and Section ?? concludes.

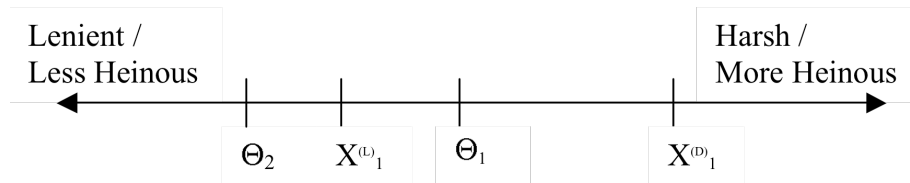
2 Model

Our model is based upon an item response model that has more generally been implemented in political science to analyze roll call data. (Jackman 2001; Martin & Quinn 2002). There is a long literature in educational testing regarding the frequentist implementation of this model, where the proto-typical example of their use is calibrating questions for the SAT:

item-response models estimate how difficult each question is, and how well each question differentiates between test takers of differing abilities. More recent Bayesian applications have been used in political science to estimate ideology of legislators and jurists. (Jackman 2001; Martin & Quinn 2002) In our application, the model measures the culpability of a given crime (how "bad" the crime is), along with individual preferences of jurists toward punishment (lenient to harsh). At the same time, the models also provide an estimate of how much individual variation in preferences matters in a given case.

The model posits a unidimensional space that determines the culpability of a defendant-crime combination (for brevity, we'll call this a "case", although case, as a legal term, may include more than one defendant-crime combination). Specifically, the culpability of an individual case is placed on a scale such as that shown in Figure 2. Here, $x^{(d)}$ and $x^{(l)}$ represent the policy outcomes for a single case if the defendant is put to death ($x^{(d)}$), or given life without parole (LWOP) ($x^{(l)}$). Similarly, each person who might judge a crime - think a jury person here - can also be represented on this scale. In Figure 2, individual i is represented by θ_i .

Figure 1: **Spatial Model of Culpability**



This model does not assume that individuals represented by θ_i discuss the case as juries would; in fact, it assumes that each person votes her own preference independently of all other participants. The set of individuals' preferences, $\{\theta_i\}$, is conceptualized as a random sample of potential participant preferences regarding the cases, and therefore represents the broader universe of preferences regarding similar cases more generally. In turn, the set of dual policy outcomes, $\{x^{(l)}, x^{(d)}\}$, combine to represent the average belief, across participants, regarding the specific cases. We reparameterize the case-specific effects, $\{x^{(l)}, x^{(d)}\}$, below into more intuitive parameters.

2.1 Formal Model

To formalize the model, let $i = 1, \dots, N$ index individuals who vote on each case, and $k = 1, \dots, K$ index the cases. Each individual i has a preferred policy position (on a continuum) with respect to harshness of punishment denoted θ_i . For each case, there are two possible policy outcomes: death (denoted $x^{(d)}$) and life without parole ($x^{(l)}$). Person i chooses between the two outcomes on the basis of a latent utility function. Person i 's utility is based upon a quadratic loss function based on the distance between person i 's preferences, θ_i and the two possible outcomes in the case, $x^{(l)}$ and $x^{(d)}$. More specifically,

$$\begin{aligned} u_{ik}^{(d)} &= -(x_k^{(d)} - \theta_i)^2 + \xi_{ik}^{(d)}; \text{ person } i \text{ votes for death in case } k \\ u_{ik}^{(l)} &= -(x_k^{(l)} - \theta_i)^2 + \xi_{ik}^{(l)}; \text{ person } i \text{ votes for LWOP in case } k \end{aligned}$$

where $\xi_{ik}^{(l)}$ and $\xi_{ik}^{(d)}$ are the person-case specific random utility shocks associated with punishment of LWOP and death, respectively. Putting these together, the overall change in person i 's utility is :

$$\begin{aligned} v_{ik}^* &= u_{ik}^{(d)} - u_{ik}^{(l)} \\ &= -(x_k^{(d)} - \theta_i)^2 + \xi_{ik}^{(d)} + (x_k^{(l)} - \theta_i)^2 - \xi_{ik}^{(l)} \\ &= -[x_k^{(d)} x_k^{(d)} - x_k^{(l)} x_k^{(l)}] + 2[x_k^{(l)} - x_k^{(d)}]\theta_i + [\xi_{ik}^{(d)} - \xi_{ik}^{(l)}] \\ &= \alpha_k + \beta_k \theta_i + \epsilon_{ik} \end{aligned}$$

where (α_k, β_k) is the vector that describes the culpability measure of a crime, θ_i is the preferred policy position of person i and ϵ_{ik} is a random shock to utility.

As is standard, we define the voting behavior of each person i in each case k in the following way:

$$v_{ik} = \begin{cases} 1, & v_{ik}^* > 0 \\ 0, & v_{ik}^* \leq 0 \end{cases}$$

Note that the v_{ik}^* are latent (unobserved) variables and the unobserved parameters are $\alpha_k, \beta_k, \theta_i$. If θ_i were observed data, this model would be a straightforward probit model under the assumption that the random shocks to utility, $\epsilon_{ik} \sim N(0, 1)$. This assumption would also fix the scale of the parameters in a probit model, which is otherwise undefined. (Note that the entire equation $v_{ik}^* = \alpha_k + \beta_k \theta_i + \epsilon_{ik}$ could be scaled by a factor of 2 and the likelihood of the data, which depends only on whether the latent variable v_{ik}^* is positive or negative, does not change).

In this case, however, the $\{\theta_i\}$ are unobserved, and the model, in frequentist terms, cannot be estimated as a probit model. From a Bayesian perspective, however, unobserved variables are easily taken care of by incorporating them into the model as parameters; that is the strategy of this paper as well. Because the $\{\theta_i\}$ are unobserved, however, this leads to two remaining identification problems that are well known in the literature on item-response models. The first is a scale problem: the relative scales of $\{\beta_k\}$ and $\{\theta_i\}$ are not identified, as one could multiply all $\{\beta_k\}$ by a factor of A and multiply all $\{\theta_i\}$ by a factor of $1/A$ without changing the likelihood of the data, even if the general scale of v_{ik}^* is fixed. In addition, the rotation of the model is not identified; multiplying both $\{\theta_i\}$ and $\{\beta_k\}$ by -1 does not change the likelihood. In a unidimensional context, this is the equivalent of flipping the positive and negative scale; it is impossible to tell from the data whether positive numbers are more lenient, or harsher. One typically defines the scale and orientation of the model fixing two θ_i , or, in Bayesian setting, providing a highly informative prior on two θ_i . Because of the structure of our model, we fix $\beta_k > 0$ to determine the orientation of the model; unlike many situations, it is clear that a vote for the death penalty is always the harsher choice. We therefore only need to fix one θ_i to fully identify the model.

Overall, this model creates a likelihood of the data and latent variables $p(v_{ik}, v_{ik}^* | \alpha_k, \beta_k, \theta_i)$ which is truncated normal:

$$p(v_{ik}, v_{ik}^* | \alpha_k, \beta_k, \theta_i) \propto \exp\{-0.5(v_{ik}^* - \alpha_k - \beta_k * \theta_i)^2\} v_{ik} \times 1_{\{v_{ik}^* > 0\}} \\ + \exp\{-0.5(v_{ik}^* - \alpha_k - \beta_k * \theta_i)^2\} (1 - v_{ik}) \times 1_{\{v_{ik}^* \leq 0\}}$$

2.2 Priors for unobservables

Because we posit a Bayesian solution to the item-response model, we need to complete the model by enumerating the priors for the unobserved parameters and latent variables. We require priors for the set of parameters $\{\alpha_k, \beta_k, \theta_i\}$ and any hyperparameters we introduce. As an initial model, we use standard normal conjugate priors for these parameters, centered at zero, with one θ_i having a very informative prior to solve the identification issue.⁴ Future work will include more flexible priors that incorporate more of the specific structure of death penalty views.⁵ We posit a inverse-gamma conjugate hyperpriors for the variances of our parameters $\{\alpha_k, \beta_k, \theta_i\}$.

The entire prior is defines as:

$$\beta_k \sim N(0, \sigma_\beta^2) 1_{\beta_k > 0} \\ \alpha_k \sim N(0, \sigma_\alpha^2) \\ \theta_k \sim N(0, \sigma_\theta^2), k \neq 1 \\ \theta_1 \sim N(1, 0.05)$$

⁴As mentioned above, one is required to fix, or identify through the prior, $(D + 1)$ θ_i s, where D is the dimensionality of the problem. We currently assume a dimensionality of $D = 1$, although we plan on expanding this to two dimensions in future iterations of this project, based upon the gruesomeness and pre-planning of the crime

⁵In our model, there is significant reason to believe that a simple normal prior does not model reality well. Specifically, recall that our participants are being asked about a highly polarizing issue: the death penalty. There are many people are categorically opposed to the death penalty. Indeed, we ask about this in the survey. If an individual consistently votes against the death penalty in all cases, the model cannot point identify their preference point; in theory, the best the model can do is to estimate the interval $(-\infty, \theta^*)$ where θ^* is less than all possible cases. In practice, the model cannot estimate probability of voting for the death penalty of zero; this is undefined. Thus, the person will be dropped from the estimation. However, if we truncate the possible values of θ and allow for a multimodal prior, we could more realistically estimate the model. Following this logic, there are also some individuals who are very likely to vote for the death penalty, although they likely recognize that some homicide cases do not warrant such a penalty (96 of the participants voted pro-death penalty in every case). These people would have to find an unusual case in order to vote against the death penalty. Together, this suggests that a mixture of three normal distributions, truncated, can accurately reflect the prior knowledge of θ : a high precision, large negative mean normal, a diffuse normal (for the middle people who don't have strong views in general) and a reasonably high precision, large positive mean normal. Identification would require truncation of the potential θ_i in this model.

$$\sigma_\beta^2, \sigma_\alpha^2, \sigma_\theta^2 \sim IG(A, B)$$

Combining the prior and the likelihood via Bayes law to determine the posterior distribution, we find a posterior of:

$$\begin{aligned} p(\beta, \alpha, \theta, V^*, \sigma_\beta^2, \sigma_\alpha^2, \sigma_\theta^2 | \{V_{ik}\}) &\propto \prod_{i,k} p(V_{ik} | V_{ik}^*) \times \prod_{i,k} p(V_{ik}^* | \beta_k, \theta_i, \alpha_k) \\ &\times \prod_k (p(\beta_k) \times p(\alpha_k)) \times \prod_i p(\theta_i) \times p(\sigma_\alpha, \sigma_\theta, \sigma_\beta) \\ &\propto \prod_{i,k} (v_{ik} \times 1_{\{v_{ik}^* > 0\}} + (1 - v_{ik}) \times 1_{\{v_{ik}^* \leq 0\}}) \\ &\times \prod_{i,k} \exp\left\{-\frac{1}{2}(v_{ik}^* - \alpha_k - \beta_k \theta_i)^2\right\} \\ &\times \prod_k \sigma_\alpha \exp\left\{-\frac{1}{2\sigma_\alpha^2}(\alpha_k)^2\right\} \times \prod_k \sigma_\beta \exp\left\{-\frac{1}{2\sigma_\beta^2}(\beta_k)^2\right\} \\ &\times \prod_i \sigma_\theta \exp\left\{-\frac{1}{2\sigma_\theta^2}(\theta_i)^2\right\} \\ &\times \frac{1}{\sigma_\alpha^{\nu-2}} \exp\left\{-\frac{\nu}{2\sigma_\alpha^2}\right\} \times \frac{1}{\sigma_\beta^{\nu-2}} \exp\left\{-\frac{\nu}{2\sigma_\beta^2}\right\} \times \frac{1}{\sigma_\theta^{\nu-2}} \exp\left\{-\frac{\nu}{2\sigma_\theta^2}\right\} \end{aligned}$$

3 Data

Our data comes from two sources. First, we gather information regarding intentional homicide cases in Missouri where prosecution was initiated between January 1, 1997 and December 31, 2001 and resulted in a homicide conviction. We gathered data on the crimes and the defendants from a variety of sources, including police investigative reports, FBI records of criminal histories, court records, newspaper articles and appellate decisions. The goal is to recreate as closely as possible the data available to the prosecutor at the time the prosecutor makes initial charging and plea bargaining decisions. We will also use this information about the cases to correlate various factors with culpability, and to control for culpability in analyses of prosecutorial decision-making in the case study below.

We use this information to create a one paragraph detailed description of each case,

which includes information about the defendant (age, sex, criminal history), and specifics about the crime. We then use these descriptions in a survey which asks participants to read the paragraph description and answer some questions about the case. First, participants are asked whether the case involved significant planning (5 point scale) and whether it was a particularly gruesome way for the victim(s) to die (5 point scale). Then, participants are asked to rank the case in three different ways. The first ranking on a 5 point scale, from least bad to worst. Second, participants are asked to vote yes or no on whether the defendant deserved the death penalty in the particular case. Finally, the participants are asked for their preferred sentence, out of 7 potential choices ranging from no prison to LWOP. There are 243 different cases;⁶ the each participant answers questions regarding a random subset of 10 cases. Using these three rankings, the survey provides three different ways to measure culpability levels. The primary method is to translate the death penalty votes into a culpability scale using the item-response model described above. A more naive method involves simply taking the average percentage of death votes for each case. However, the survey provides two other more direct measurements of the culpability of each case: the answer to the scale question (how heinous was the crime), and the answer to the sentencing question. Finally, as a robustness check, the answers to these second two scalings can be grouped into dichotomous variables that would allow for the same item-response model to be estimated using these alternative specifications.

We implemented the survey through a survey research group; they sent invitations to $m=604$ individuals. The current results are based upon $n=531$ individual responses (data collected is ongoing, and has reached over 90% participation). Each individual was paid \$6 to complete the study; most took 15-25 minutes to do so. Three individuals were excluded because they were not U.S. citizens.

One additional data issue results from our substantive issue of the death penalty. A nontrivial percentage of individuals are against the death penalty under any circumstances; these people will vote against the death penalty for all crimes, without thought to the culpa-

⁶These cases are a stratified random sample of all 1046 homicides. We stratified the original sample on whether the death penalty was sought in the case, including all such cases (128) and a random sample of the remaining 918 cases.

bility of the particular crime at issue. In the general population, between 20-30% of people identify as anti-death penalty in opinion polls. For any individual who votes against the death penalty in every case they see - a superset of the set of people categorically against the death penalty - the model cannot identify θ_i because there is no variation in the dependent variable. Similarly, there may be individuals who always, or almost always, would vote for the death penalty in a homicide; if there is no variation in the 10 random cases that our participants review, the model cannot estimate an ideal point for these individuals. These individuals, however, also provide no information about the relative culpability of the cases in our database, and so dropping these individuals from the sample is unproblematic. We drop 175 individuals who always voted for or always against the death penalty. Similarly, cases upon which all participants agree are uninformative, and must be dropped for the estimation to converge; we drop one case in which all individuals agreed that the individual should not get the death penalty. As mentioned, dropping this first level of data is not problematic, because the data do not provide information. The problem becomes that dropping the individuals who always vote against the death penalty lowers the variance of high-culpability crimes, making it more likely that all of the remaining participants in the study voted for the death penalty. The same logic applies to people: dropping the cases on which everyone agrees makes it possible that some individuals have now (in the smaller sample) only voted in one direction. We iterate this approach, dropping individuals or cases with no variance in the dependent variable, until every case and person displays variation in the dependent variable. These are individuals that are not categorically for or against the death penalty, and cases that are not so clear that everyone (in the full sample) would vote the same way.

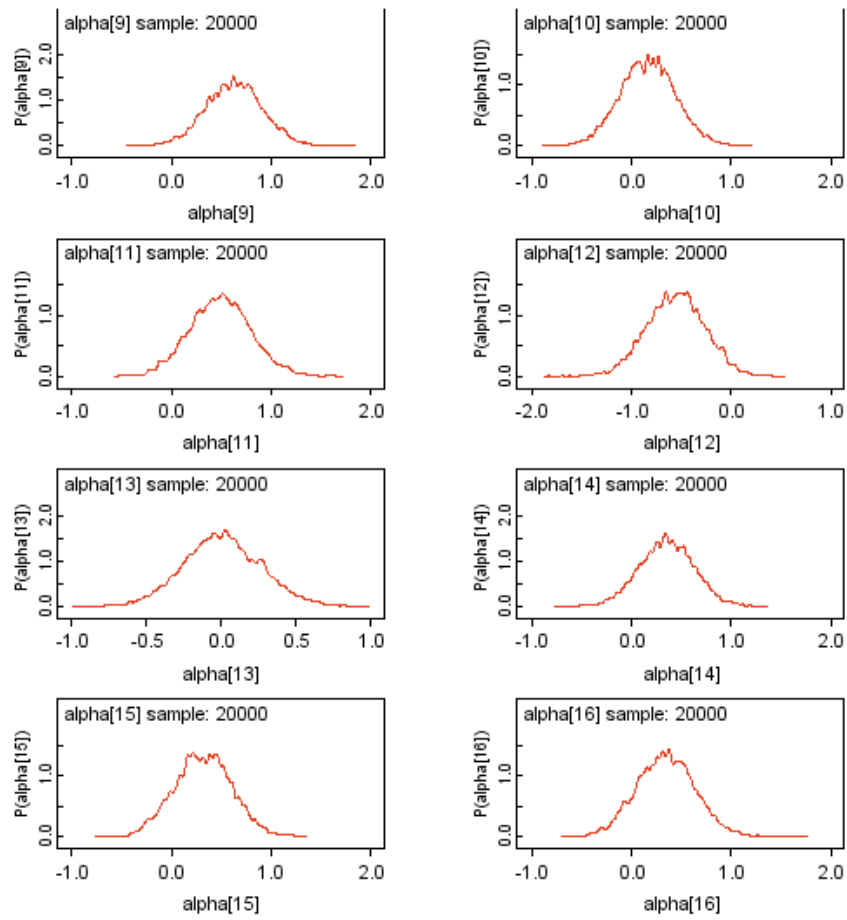
By dropping these individuals and cases from the sample, we do lose some information, therefore. We are, however, forced to drop them from the data absent further modeling of the specific polarizing nature of the death penalty. While not ideal, three specifics suggest that this should not be too troubling. First, not very many people or cases are dropped after the initial round (which contain no information). We drop nine individuals (out of a remaining 356 individuals who are not completely polarized on the issue), and four out

of 242 remaining cases. Second, these are the cases and individuals who had very strong preferences - and lower variance of the dependent variable - in the first place. They therefore provide less information than a similar individual or case in which there was more disagreement. Finally, with respect to the case parameters, we do not require a random sample of individuals to accurately measure the case parameters; we require only the assumption that the subsample of individuals we have are using the same general rubric to determine culpability as the larger random sample of individuals (the particular scale may be different, but the relative positions will be the same). In all, we have 347 individuals who collectively answering questions about 238 cases; because a small number of individuals did not answer one or two questions about a given case, we have a sample size of $n=3339$ person-case data points.

4 Initial Results: Identification of Parameters

The initial results for the model are promising; they demonstrate that the model can identify between cases and individual subjects. The model converges quickly; the results presented here are based upon 10,000 draws after an initial burn-in of 1000 draws. Auto-correlation is minimal as well. Initially, the results demonstrate that the parameters θ , β , and α are identified in the model.

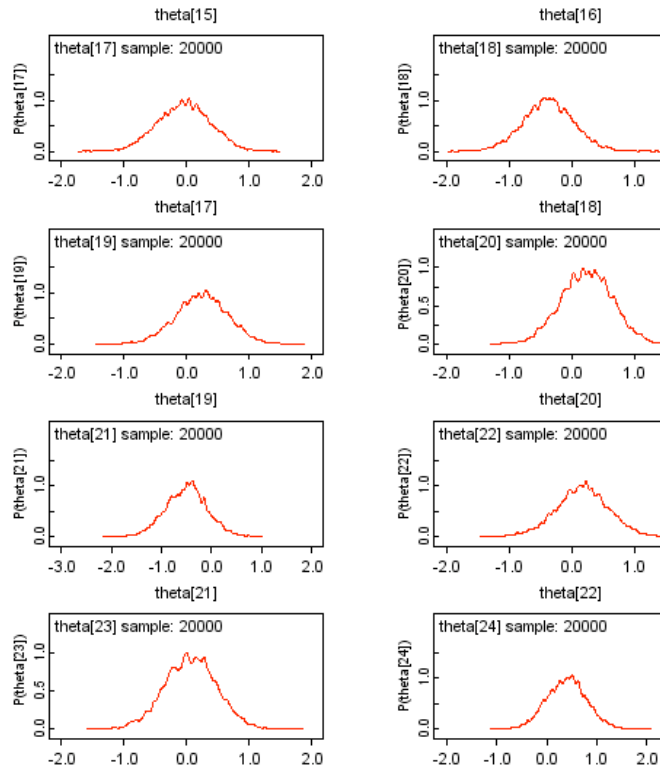
Figure 2: Density of Select α_k



More specifically, it is possible to distinguish between different cases; that is, the data are measuring that cases have different levels of culpability, and participants have different ideal points for each case. For example, Figure 2 shows a subset of the distributions for the α parameters. The means of α distributions range from $\alpha = 0.85$ to $\alpha = -0.69$. It is these means that we use to measure the average culpability of each crime in our case study below.

The θ parameters, which describe how each person responds, on average, to cases, are similarly disbursed. Figure 3 provides a graph of the distribution of the θ parameters. While there is considerable overlap in the distributions across people, they remain distinct

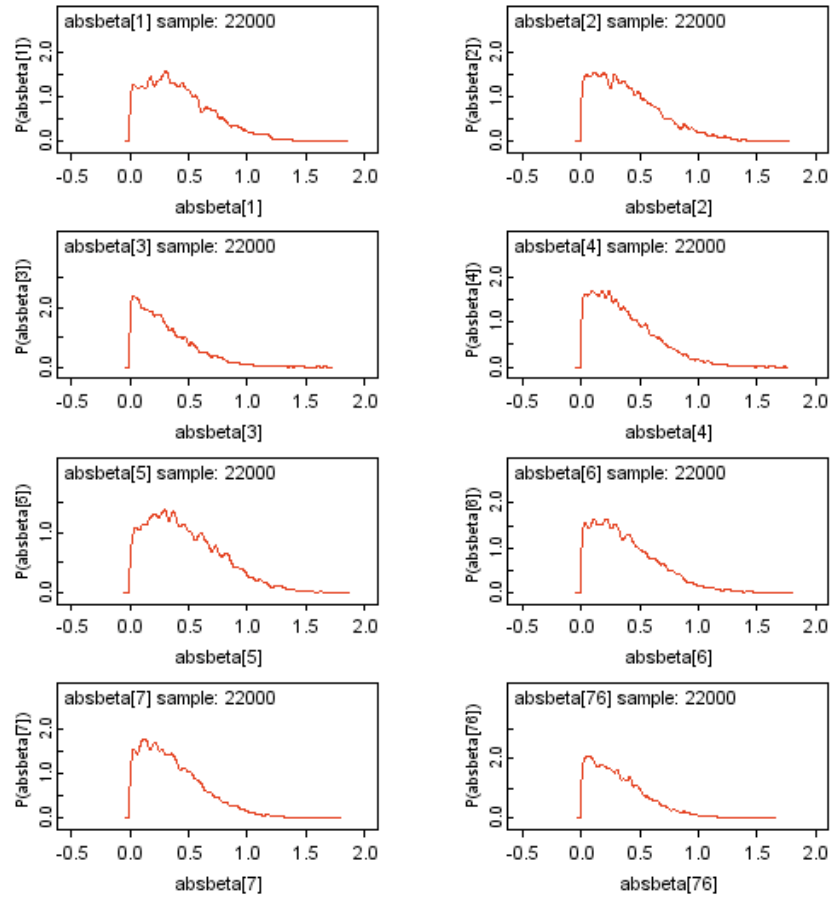
Figure 3: Density of Select θ_k



distributions, with significantly different means and variances. The means for individuals in our sample range from $\theta = 0.48$ to $\theta = -0.69$. Finally, Figure 4 demonstrates that

the β parameters are more similar to each other, but also demonstrate differently shaped distributions; some are more censored than others.

Figure 4: Density of Select β_k



5 Case Study: Death Penalty in Missouri

In this section, we demonstrate the use of our measures to determine whether the death penalty is applied in a biased manner in Missouri. Since the time of David Baldus landmark study of the Georgia death penalty, one of the most consistent findings in the rich body of prior research on the death penalty is that race of the participants and place of the crime matters. (Baldus et al. 1987; Baldus et al. 2003). In particular, white victims are afforded higher status, and, in some cases, that black defendants are treated more harshly, particularly if their victims are white. In addition, urban and rural areas treat death penalty

prosecution differently, although though there is no consistent direction of the difference (sometimes urban communities are harsher; sometimes rural communities are).

In prior research on the death penalty in Missouri, we found a significant pattern of geographic disparities, and, perhaps due to the geographic disparities, some disparities in the manner in which different racial groups were treated, both as defendants and when victims. (Barnes, et al. 2009). While this general finding is consistent with the prior literature, we did not control for culpability or perform any causal analysis. This example fills that gap. We investigate four decision points in the chain of events that lead to a death verdict: the initial M1 charge, a requirement for a death verdict;⁷ the M1 conviction; the decision to pursue the death penalty (made pre-trial); and the final verdict of the death penalty.

At this early stage, we investigate two possible controls for culpability: the mean of the α_k distribution for each case k , and the percentage of individuals who voted for the death penalty in each case within our survey. This second control is a more naive version of the mean α_k control, as it does not incorporate the effects of a different subset of individuals rating the case. As the cases are assigned randomly, however, this does not bias the measure, but instead simply adds additional measurement error to the estimate.

Table 1: Initial Charge of M1

	No Controls	Naive Control	α -measure Control
White D/ Black V	∞	∞	∞
Black D/ White V	1.1	0.9	0.9
Black D/Black V	1.0	1.0	1.0
Other-race V or D	0.5	0.3	0.2
SL City	2.0	1.6	1.7
Jackson County	0.3*	0.2*	0.2*
MSA, 0–10% Minority	0.6	0.3	0.3
Rural, 5–30% Minority	0.9	0.6	0.5
Rural, 0–5% Minority	1	0.7	0.7
Culpability	–	291.4***	19.2***

⁷our prior research argues that there is no substantive difference between M1 and M2 in the Missouri statute, while the penalties are significantly different: the *minimum* penalty for M1 is LWOP, while the *maximum* penalty for M2 is life with the possibility of parole.

Table 1 provides the results for three logistic regression models describing the relationship between the M1 charging decision, the race of defendant and victim, and the geographic region of the crime. This last variable is broken down into 6 categories. Each county in Missouri has a separate county prosecutor, who makes independent decisions regarding charging and plea bargaining behavior. Because most counties have very few homicides during the five years of our study, we grouped suburban and rural counties by a MSA (metropolitan statistical area) or rural designation, as well as a percentage of the jury population that is not white (high or low, essentially). Saint Louis City and Jackson County (Kansas City) are each a separate category.

The baseline value in Table 1 and the remaining tables in this section is a crime committed by a white defendant, against a white victim, in an MSA county with reasonably high minority jury pool (10–30%)(generally the suburbs of St. Louis City and Kansas City, along with the county of the state capital). The second column lists the odds ratio of a logistic regression without any controls for culpability; the primary finding is that Jackson County charges M1 at about one third the rate of the baseline. Note that all white defendants who killed black victims were charged with M1; however, as there were only 4 of these cases, this finding isn't particularly robust. (All other groups have at least 20 cases each).

Controlling for culpability does not significantly change this finding, nor does it drastically change other parameters. Both the naive and α - measures provide similar results as well: Jackson County charges M1 at about one fifth the rate of the baseline. In theory, differences in charging cases should be based upon culpability of the crime alone; while culpability is the most important factor in the decision to charge M1, the geographic region of the crime is also significant.⁸

Table 2 provides analogous results investigating the final conviction decision, from the universe of individuals who could have been charged with M1, to those who were convicted

⁸Culpability is the most important individual predictor of the dependent variable in each of the four tables presented here. It is important to note, however, that culpability is not always related to decisions made during the homicide prosecution. Specifically, in logistic regressions involving decisions made by juries - whether to convict after a trial, and whether to give the death penalty after a penalty phase - the culpability measures are not statistically significant.

Table 2: M1 Conviction

	No Controls	Naive Control	α -measure Control
White D/ Black V	0.4	0.5	0.3
Black D/ White V	2.0	1.8	2.3
Black D/Black V	1.2	1.3	2.0
Other-race V or D	0.3	0.2*	0.5
SL City	0.4	0.2*	0.1**
Jackson County	0.3*	0.2*	0.4
MSA, 0–10% Minority	0.4	0.2	0.5
Rural, 5–30% Minority	0.7	0.4	0.4
Rural, 0–5% Minority	1.1	0.7	0.9
Culpability	–	762.5***	21.2***

of M1. For this to happen, the prosecutor could not have withdrawn the M1 charge due to a plea bargain, and either obtained a favorable M1 plea bargain or tried the defendant, with the jury returning a verdict of guilty on the M1 charge. Here, without any culpability controls, only crimes committed in Jackson County are significantly different from baseline; again, defendants in Jackson County are about one third as likely to be convicted of M1 as their counterparts in the suburbs. (This is due almost exclusively to the initial difference in charging rates). However, here we do see significant changes when controlling for culpability. Using the naive control, Jackson County conviction rates remain significantly lower than the baseline, but St Louis City also has significantly lower M1 conviction rates. Finally, cases with a defendant or victim who is neither Black nor white also were treated more leniently after controlling for culpability.

Under the more robust culpability measure, however, this last finding disappears, as does the significant difference seen in Jackson County. Which culpability measure is better is, perhaps, a matter of opinion. As mentioned above, the naive culpability measure incorporates more measurement error, because it is based solely on the 10 individuals who answered questions about a specific case. Ten individuals is not a large sample upon which to base an estimate. The naive measure does not, however, require one to agree with the underlying utility model that forms the basis of the α -measure. Thus, if one is skeptical about the underlying model, the naive measure is preferred; if one finds the underlying utility model reasonable, that the α -measure has the advantage that it incorporates all of the

data into the model, rather than just a small subset of the data. This allows for significantly greater precision in the estimates. The two measures are, however, substantially similar.

What remains clear, however, is that in Saint Louis City, defendants who commit similarly culpable crimes are significantly less likely to be convicted of M1 than those individuals who do so in the suburbs, such as Saint Louis County. The different average culpability levels of the crimes in these two geographic regions masked this difference in the no controls model. These differences have serious repercussions: the difference between a M1 conviction and a M2 conviction is the difference, on average, of over 30 years. M2 convictions carry a penalty of 15 years to life (with parole possibility); M1 convictions require either the death penalty or LWOP; therefore, the ranges of possible penalties for M1 and M2 do not overlap. Defendants who travel across the city limits to stick up the Quickie Mart – and shoot the store clerk in the process - face drastically different penalties depending on which side of the city limits they are on at the time of the crime. This type of arbitrariness is one of the main concerns of Supreme Court jurisprudence regarding the Eighth Amendment and the death penalty, although the Court has not embraced expanding this concern outside of the narrow context of the death penalty and into M1 convictions more broadly.

Table 3: Initial Capital Charge

	No Controls	Naive Control	α -measure Control
White D/ Black V	1.5	1.6	0.4
Black D/ White V	0.9	0.5	1.8
Black D/Black V	0.6	0.5	1.4
Other-race V or D	0.5	0.2	0.2
SL City	0.5	0.3*	0.2*
Jackson County	0.06***	0.05***	0.2*
MSA, 0–10% Minority	0.9	0.7	0.3
Rural, 5–30% Minority	2.0	1.6	0.3
Rural, 0–5% Minority	1.4	0.9	0.7
Culpability	–	610.2***	33.8***

Investigating the prosecutor’s decision to pursue capital punishment - the initial capital charge (separate from the M1 charge) - demonstrates that place matters in more than just

M1 charging and conviction; it matters in the selection of defendants who face a capital charge, and, as Table 4 suggests, in whether a defendant is sentenced to die as well.

Table 3 provides the details. The model with no controls demonstrates a significant disparity between defendants in Jackson County and other defendants: Jackson County defendants face capital charges one seventeenth as many times as defendants in the suburbs do. Capital charges are not an insignificant event; beyond the chance that the defendant will be sentenced to death (which is still reasonably unlikely at this point), a capital trial is very different from a non-capital M1 trial, in that the jury must be "death-qualified"; research has found that this makes the jury both more likely to impose the death penalty and more likely to convict than a standard jury. Capital charges, along with initial M1 charges, are also powerful tools for coercing favorable pleas.

The geographic disparity is tempered, however, in the α -measure model controlling for culpability. Here, the Jackson County defendants face a one fifth the odds of facing capital charges as defendants in nearby suburbs would face. The naive control leaves the disparity essentially untouched; overall, controlling for culpability does not completely mitigate the geographic disparities. Indeed, in the case of Saint Louis City, it exacerbates them, making the estimated disparity larger and statistically significant. Regarding the capital charging decisions made by the counties, Jackson County and St. Louis City are significantly different than their suburbs, which suggests, again, arbitrary treatment on the basis of a short car ride between neighborhoods.

Finally, Table 4 displays the results of the logistic regressions of the final decision of interest: whether the defendant received the death penalty, out of all the possible intentional homicide cases in the sample. Once again, the disparities in treatment occur across geographic borders; St. Louis City, Jackson County and MSAs with small minority populations are all less likely to sentence defendants to death than rural counties or MSA counties with larger minority populations (essentially, suburbs of the major cities along with a few counties in the south eastern portion of the state, and Jefferson City, the state capital). Jackson County, for example, sentenced no one to death - out of a total of 35 cases). Controlling for culpability does not change this basic result, although the specific size of the disparity

Table 4: Verdict of Death

	No Controls	Naive Control	α -measure Control
White D/ Black V	17.4	14.1	11.7
Black D/ White V	3.7	1.5	1.6
Black D/Black V	0.6	0.5	0.4
Other-race V or D	0	0	0
SL City	0.04*	0.03*	0.4*
Jackson County	0	0	0
MSA, 0–10% Minority	0.1*	0.1**	0.1**
Rural, 5–30% Minority	1.3	1.0	0.7
Rural, 0–5% Minority	0.8	0.5	0.5
Culpability	–	1568.5***	48.6***

varies in St. Louis City across the two different culpability measures. What is particularly striking in this Table is the similarity to Table 3: for both St. Louis City and Jackson County, the initial disparity created by the prosecutor’s initial decision remains, after all the other checks and balances in the system. This provides one more suggestion why initial decisions by prosecutors are so important to the system, and need oversight.

6 Conclusion

In this paper, we posit a latent utility model of attitudes toward the death penalty that is used to create a continuous measure of culpability. The model therefore estimates the culpability of each individual case, as well as providing a rich dataset of the individual preferences of potential jurors. We use this measure, as well as others, to control for culpability in a case study of death penalty practice in Missouri. We find large, statistically significant geographic disparities in the charging and sentencing outcomes for defendants. Contrary to the claims of prosecutors in the system, these disparities cannot be explained by the culpability of the crime: controlling for culpability does little to change the observed disparities. This suggests that the decision-making system in Missouri allows improper factors to influence decisions to a substantial degree. This, in turn, implies that the system allows for unchecked arbitrariness in the application of the death penalty based solely on a

mile's drive. While we do not find large racial disparities after controlling for geographic differences, racial housing patterns suggest that geographic disparities can induce racial disparities in outcomes. In sum, the study demonstrates that the charging and sentencing practices for homicides of the criminal justice system in Missouri uses improper factors in decisionmaking, leading to arbitrary differences in defendant outcomes.

7 Appendix A

This Appendix details the Markov Chain Monte Carlo (MCMC) algorithm used to solve the model, and provides some convergence and robustness checks. [Note: this section details the iteration of the algorithm that does not incorporate a mixture of normals; the basic premise, however, is the same for both]

7.1 MCMC Algorithm

The MCMC algorithm that is used to simulate the posterior distribution of the parameters, $\beta_k, \alpha_k, \theta_i$ is a gibbs sampling program detailed below. Gibbs sampling allows for the iterative sampling of each parameter (or block of parameters), conditional on all other parameters. Here, the algorithm breaks the parameters into five groups: $\beta = \{\beta_k, k = 1, \dots, K\}$; $\alpha = \{\alpha_k, k = 1, \dots, K\}$; $\theta = \{\theta_i, i = 1, \dots, N\}$, $\Sigma = \{\sigma_\alpha^2, \sigma_\beta^2, \sigma_\theta^2\}$ and $V^* = \{v_{ik}^*, i = 1, \dots, N, k = i, \dots, K\}$. The full posterior distribution is, once again, given below. [This section needs to be updated to include the mixture of normals priors]

$$\begin{aligned}
 p(\beta, \alpha, \theta, V^*, \Sigma | \{V_{ik}\}) &\propto \prod_{i,k} p(V_{ik} | v_{ik}^*) \times \prod_{i,k} p(V_{ik}^* | \beta_k, \theta_i, \alpha_k) \\
 &\times \prod_k p(\beta_k) \times \prod_k p(\alpha_k) \times \prod_i p(\theta_i) \times p(\sigma_\alpha, \sigma_\theta, \sigma_\beta) \\
 &\propto \prod_{i,k} (v_{ik} \times 1_{\{v_{ik}^* > 0\}} + (1 - v_{ik}) \times 1_{\{v_{ik}^* \leq 0\}}) \\
 &\times \prod_{i,k} \exp\left\{-\frac{1}{2}(v_{ik}^* - \alpha_k - \beta_k \theta_i)^2\right\} \\
 &\times \prod_k \sigma_\alpha \exp\left\{-\frac{1}{2\sigma_\alpha^2}(\alpha_k)^2\right\} \times \prod_k \sigma_\beta \exp\left\{-\frac{1}{2\sigma_\beta^2}(\beta_k)^2\right\} \\
 &\times \prod_i \sigma_\theta \exp\left\{-\frac{1}{2\sigma_\theta^2}(\theta_i)^2\right\} \\
 &\times \frac{1}{\sigma_\alpha^{\nu-2}} \exp\left\{-\frac{\nu}{2\sigma_\alpha^2}\right\} \times \frac{1}{\sigma_\beta^{\nu-2}} \exp\left\{-\frac{\nu}{2\sigma_\beta^2}\right\} \times \frac{1}{\sigma_\theta^{\nu-2}} \exp\left\{-\frac{\nu}{2\sigma_\theta^2}\right\}
 \end{aligned}$$

Assuming a complete draw $\beta^{(t-1)}, \alpha^{(t-1)}, \theta^{(t-1)}, V^{*(t-1)}, \Sigma^{(t-1)}$ at iteration $(t - 1)$, iteration t is drawn as follows:

1. *Draw* $\beta^{(t)}$ The full conditional for each β_k is given by:

$$\begin{aligned} p(\beta_k | \alpha, \theta, V^*, \Sigma\{V_{ik}\}) &\propto p(V_{ik}^* | \beta_k, \theta_i, \alpha_k) \times p(\beta_k) \\ &\propto \exp\left\{-\frac{1}{2}(v_{ik}^* - \alpha_k - \beta_k \theta_i)^2\right\} \times \exp\left\{-\frac{1}{2\sigma_\beta^2}(\beta_k)^2\right\} \end{aligned}$$

where the β_k are independent conditional on $\{\alpha, \theta, V_{ik}^*\}$. Dropping the portion of the exponent that does not depend on β_k and completing the square gives the following:

$$p(\beta_k | \alpha, \theta, V^*, \Sigma, \{V_{ik}\}) \propto \exp\left\{-\frac{1}{2}\left(\frac{1}{\sigma_\beta^2} + \sum_{i=1}^N \theta_i^2\right)\left(\beta_k - \frac{\sum_{i=1}^N (v_{ik}^* - \alpha_k)\theta_i}{\frac{1}{\sigma_\beta^2} + \sum_{i=1}^N \theta_i^2}\right)^2\right\}$$

This last expression is the kernel of a Normal density with mean $\frac{\sum_{i=1}^N (v_{ik}^* - \alpha_k)\theta_i}{\frac{1}{\sigma_\beta^2} + \sum_{i=1}^N \theta_i^2}$ and variance $(\sigma_\beta^{-2} + \sum_{i=1}^N \theta_i^2)^{-1}$. Thus, we draw $\beta_k^{(t)}$ from $\beta_k | \theta^{(t)}, \alpha^{(t)}, V_{ik}^{*(t)}, \sigma_\beta^{2(t)} \sim N\left(\frac{\sum_{i=1}^N (v_{ik}^* - \alpha_k)\theta_i}{\frac{1}{\sigma_\beta^2} + \sum_{i=1}^N \theta_i^2}, (\sigma_\beta^{-2} + \sum_{i=1}^N \theta_i^2)^{-1}\right)$.

2. *Draw* $\alpha^{(t)}$ We can similarly calculate the full conditional distribution of the α_k , which are independent conditional on the other parameters. Doing so, we find that $\alpha_k | \beta, \theta, V_{ik}^*, \sigma_\alpha^2 \sim N\left(\frac{\sum_{i=1}^N (v_{ik}^* - \theta_i \beta_k)}{N + \sigma_\alpha^{-2}}, (N + \sigma_\alpha^{-2})^{-1}\right)$. Because this is the second step in the Gibbs Sampler, we fix β_k at their values from the current iteration, and all other parameters at their values from the previous iteration when drawing α_k .

3. *Draw* $\theta_i^{(t)}$

$$\text{Similarly, } \theta_i | \alpha, \beta, \sigma_\theta^2, V_{ik}^* \sim N\left(\frac{\sum_{k=1}^K \beta_k (v_{ik}^* - \alpha_k)}{\sum_{k=1}^K \beta_k^2 + \sigma_\theta^{-2}}, (\sum_{k=1}^K \beta_k^2 + \sigma_\theta^{-2})^{-1}\right)$$

4. *Draw* v_{ik}^*

The v_{ik}^* are truncated normals, conditional on all other parameters and the data.

5. *Draw* $\sigma_\alpha^2, \sigma_\beta^2, \sigma_\theta^2$

Each of the σ_j^2 are inverse χ^2 , conditional on α, β, θ . They are conditionally independent of the data and the latent variables.