

# A Bayesian Hierarchical Model for Estimating Heterogeneous Effects

Joshua Alley  
Assistant Professor  
University College Dublin\*  
joshua.alley@ucd.edu

August 8, 2023

Abstract

---

\*Thanks to Carlisle Rainey for helpful comments.

# 1 Introduction

Every independent variable social scientists study impacts some units more than other. Such heterogeneous effects are present in observational and experimental studies. As a result, how different units respond to the same stimulus is essential for policy and scholarship.

This note introduces a hierarchical Bayesian approach to estimating heterogeneous effects. The model estimates how groups respond to an key independent variable based on unit characteristics, context and other experimental treatments. Modeling heterogeneous effects in this way is easy to interpret, which facilitates argument testing. It also allows for comparisons between different sources of heterogeneous effects.

This model complements the extensive array of tools for estimating heterogeneous effects. Early attempts used parametric interactions and subgroup analyses. Latter tools draw on random forests (Green and Kern, 2012; Wager and Athey, 2018), support vector machines (Imai and Ratkovic, 2013), and ensemble methods (Grimmer, Messing and Westwood, 2017; ?).

Hierarchical modeling complements existing techniques. Simple interactions and subgroup analysis are easy to interpret may understate variation by only examining a few factors, and do not scale well across more than a few dimensions. More complex machine learning techniques capture complex patterns, but can be harder to interpret. Using a hierarchical model where other variables predict the main effect a research is interested in preserves a simple and interpretable structure like interactions, while accommodating more factors and reducing the risks of underpowered subgroup analysis. This facilitates argument testing, at the cost of flexibility to discover high-dimensional heterogeneity.

To demonstrate how this model works, I reanalyze a study of...

## 2 A Hierarchical Model of Heterogeneous Effects

The model uses at least two equations. The first equation links the treatment and outcome. The second equation estimates heterogeneous effects for each unique combination of variables that could modify the impact of a key independent variable.<sup>1</sup>

For the outcome equation, start with  $N$  units indexed with  $i$ , some of which receive a binary experimental treatment  $T$ .<sup>2</sup> For simplicity, I describe a model with an outcome variable  $y$  that is normally distributed with mean  $\mu_i$  and variance  $\sigma$ . This approach can apply to binary, categorical and other outcomes as well.

The first equation is a simple random effects regression. The outcome for each unit is then a function of an overall intercept  $\alpha$ , a matrix of control variables  $\mathbf{X}$ ,<sup>3</sup> and a set of varying intercepts  $\theta$ , which are normally distributed with mean  $\eta$  and variance  $\sigma_\theta$ . All units are divided into  $g$  groups based on unique combinations of treatment status and predictors of heterogeneous effects  $\mathbf{Z}$ . Each  $\theta$  parameter estimates how much a specific group deviates from the overall intercept  $\alpha$  as a result of the treatment.

$$\begin{aligned}
 y &\sim N(\mu_i, \sigma) && \text{(Likelihood)} \\
 \mu_i &= \alpha + \theta_g + \mathbf{X} * \beta && \text{(Outcome)} \\
 \theta_g &\sim N(\eta_g, \sigma_\theta) \\
 \eta_g &= \alpha_g + \mathbf{Z} * \lambda && \text{(Heterogeneous Effects)}
 \end{aligned} \tag{1}$$

The second equation then predicts the group intercepts.  $\mathbf{Z}$  can contain unit characteristics, other treatments, or contextual factors; anything that modifies the impact of the treatment. The second equation also includes an intercept  $\alpha_g$  that is the impact of a treatment when all the heterogeneous effect variables are equal to zero. As in other regressions, this intercept may

---

<sup>1</sup>If a researcher wants to estimate heterogeneous effects for multiple variables, they can add additional heterogeneous effect equations.

<sup>2</sup>I describe an extension to continuous interventions in the appendix.

<sup>3</sup>This can be omitted, depending on the application.

not always be substantively meaningful.

Identification requires one restriction. One group is the control group, and I fix that  $\eta$  to zero and set the corresponding predictors of that parameter to zero as well. Without this restriction, sampling from the full posterior is difficult at best.

As a result of the identification restriction, the control group is equal to the overall intercept, and non-zero  $\theta$  parameters estimate how much a given treated group deviates from that mean. The model thus estimates heterogeneous effects by group, where each group has a unique mix of factors that can generate heterogeneous effects. Partial pooling means that groups with fewer units still have reasonable uncertainty estimates.

The  $\theta$  and  $\lambda$  parameters are the key estimates for understanding heterogeneous effects. Variables with a positive  $\lambda$  parameter increase the impact of treatment, while a negative  $\lambda$  indicates that a factor reduces the impact of the treatment.<sup>4</sup> The net impact of treatment on a group  $\theta$  depends on the balance of these factors.

Estimating heterogeneous effects in this way has three advantages. First, this model allows researchers to account for multiple potential sources of heterogeneous effects in an easy to interpret framework. Researchers can thus examine theories of heterogeneous effects and compare sources of variation. Partial pooling also facilitates reasonable estimates for small groups by sharing information across treated units and leveraging variation in the predictors of heterogeneous effects. Finally, this approach will be faster than machine learning approaches for many datasets, though it may not scale well to big datasets.

Like all methods, this approach has downsides as well. First, it is most straightforward for discrete treatments and discrete modifiers of those treatments. Using continuous variables in the heterogeneous effects equation creates many small groups or individual treatment effects, which increases the risk of sampling problems. Researchers can create bins of continuous vari-

---

<sup>4</sup>Whether the  $\lambda$  parameters have a direct substantive meaning depends on the likelihood and whether the researcher uses a link function.

ables, but reducing variation in this way could impact inferences. Second, this approach will not capture complex, high-dimensional interactions, which machine learning approaches are better placed to find. That said, researchers can inject substantial flexibility into this modeling approach, using interactions or non-linear specifications in either level of the model. Third, this model can show general trends, but will not be able to make powerful comparisons between all pairs of groups.

### **3 Application**

### **4 Conclusion**

## References

- Green, Donald P and Holger L Kern. 2012. “Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees.” *Public Opinion Quarterly* 76(3):491–511.
- Grimmer, Justin, Solomon Messing and Sean J Westwood. 2017. “Estimating Heterogeneous Treatment Effects and the Effects of Heterogeneous Treatments with Ensemble Methods.” *Political Analysis* 25(4):413–434.
- Imai, Kosuke and Marc Ratkovic. 2013. “Estimating treatment effect heterogeneity in randomized program evaluation.” *The Annals of Applied Statistics* 7(1):443–470.
- Wager, Stefan and Susan Athey. 2018. “Estimation and Inference of Heterogeneous Treatment Effects using Random Forests.” *Journal of the American Statistical Association* 113(523):1228–1242.