CHILD WELFARE PREDICTIVE RISK MODELS AND LEGAL DECISION MAKING

Author:

Matthew Trail, Research Fellow with the Max Planck Institute for Research on Collective Goods

Email:

trail@coll.mpg.de

Researcher Page:

https://www.coll.mpg.de/matthew-trail.html

OSF Page:

https://osf.io/emcp3/

CHILD WELFARE PREDICTIVE RISK MODELS AND LEGAL DECISION MAKING

ABSTRACT

Background: Child welfare agencies around the world have experimented with algorithmic predictive modeling as a method to assist in decision making regarding foster child risk, removal and placement.

Objective: Thus far, all of the predictive risk models have been confined to the employees of the various child welfare agencies at the early removal stages and none have been used by attorneys in legal arguments or by judges in making child welfare legal decisions. This study will show the effects of a predictive model on legal decision making within a child welfare context.

Participants and Setting: Lawyers, judges and law students with experience in child welfare or juvenile law were recruited to take an online randomized vigentter survey.

Methods: The survey consisted of two vignettes describing complex foster child removal and placement legal decisions where participants were exposed to one of three randomized predictive risk model scores. They were then asked follow up questions regarding their decisions to see if the risk models changed their answers.

Results: Using structural equation modeling, high predictive model risk scores showed consistent ability to change legal decisions about removal and placement across both vignettes. Medium and low scores, though less consistent also significantly influenced legal decision making.

Conclusions: Child welfare legal decision making can be affected by the use of a predictive risk model, which has implications for the development and use of these models as well as legal education for attorneys and judges in the field.

Keywords: predictive risk models, legal decision making, child welfare/foster care,

INTRODUCTION

Predictive models have long had supporters in the child welfare world as they potentially represent a way for state Child Protective Service (CPS) agencies to make better decisions and protect more children from harm. However, the models have a mixed history in the child welfare world.

Despite their widespread use and history, none of the predictive models have ever been used by attorneys representing parties in a CPS case or by the courts that are hearing the cases. However, courts in other areas of the law, such as criminal recidivism have experience using predictive risk models (Hamilton, 2019), so the likelihood remains that eventually these models will make it to the courtroom even if only as discoverable evidence presented by a party.

How these models might affect legal decisions amongst actual practitioners in child welfare is unknown. The present study is the first of its kind to test how legal decision making with the use of predictive risk models in the child welfare context. In an experimental survey of attorneys, law students and judges with child welfare experience, the results show that predictive risk models can change legal decisions about the placement and removal of foster children. This has important ethical and policy issues that will be discussed later in this paper.

Research has supported the finding that machine advice can affect decision making (Bogert, et al., 2021; Grgić-Hlača, et al., 2022) and that humans are ambivalent about the machine advice given (Burton et al., 2019). There are also serious critiques of the use of predictive models across fields especially in the light of their tendency to amplify racial bias and discrimination such as within hiring decisions (Raghaven et al., 2019); medical decisions (Obermeyer, et al., 2019), facial recognition (Raji & Fried, 2021), and policing (Collins, 2018) to name just a few examples.

Though this paper could not hope to trace the use of predictive models across fields, nor even to identify each and every predictive model used in a child welfare context, a brief highlighted history of their use in the child welfare system to the present day as well as a brief discussion of how courts and attorneys encounter algorithms in other legal areas will hopefully help put this current research in the proper context.

LITERATURE REVIEW

Samant and colleagues (2021) with the American Civil Liberties Union (ACLU) reported that as many as 26 states had experimented with child welfare predictive models, and at least 11 were currently using them, though this this number might be higher as it is not always clear when and where the models are being used. The ACLU noted that most of the models are used for some sort of child abuse risk modeling, though those models could come in different forms such as individual risk models for individual children or even neighborhood risk models for all children in a certain location (Samant, et al., 2021). Researchers who examined one model out of Washington state found that the model being used did not reflect many subjective factors that caseworkers were considering and the factors it was using were not significantly correlated with abuse (Saxena, et al., 2022).

Though as described above, most models are focused on risk and removal, other uses of predictive models in child welfare include their use to predict permanency outcomes of youth already in care (Ahn, et al., 2021; Stepura, et al., 2021; Elgin, 2018), to predict the stability of reunification (Purdy & Glass, 2020) to predict the most stable placement (Moore, et al., 2016), to predict which youth are likely to run away from placement (Chor, et al, 2022), and which cases could best be served by a team of child welfare experts (Willcott & Stewart, 2021).

Leslie and colleagues (2020) stated that the use of predictive models in child welfare was predicated on the possible benefits of the models such as more consistent and rational decision making, better CPS accountability and better resource management. Drake et al. (2020) noted that the use of actuarial tools for prediction in child welfare is very common and even those tools require human input that can be flawed. They also pointed out that increased accuracy in decision making comes with its own ethical impetus as these tools potentially have the power to impact the lives of real humans (Drake, et al., 2020).

One of the common complaints levied against the use of predictive models is that the models are biased. Gillingham (2020) noted that the biases can be introduced both through the absence of the correct data within the case files from which it is based and also the biases inherent in human decisions that are reflected within the data. Keddell (2019) discussed that the child welfare system includes both parents and children and that sometimes those parties have conflicting interests and conflicting ideas of fairness.

Perhaps the most discussed of the current models is the Allegheny Family Screening Tool (AFST) used in Allegheny County, Pennsylvania, the county that encompasses Pittsburgh. Since the AFST is well known in the literature and models based on it have expanded to new jurisdictions, I chose it as the inspiration for the hypothetical model used in my experimental survey. I will discuss its history, purpose, controversy and current use at more length.

In the report of the origin of the AFST, Vaithianathan, et al. (2017) described how predictive modeling had been used already in medical settings and that the AFST was created following a 2014 request from Allegheny County CPS. The county is special because it maintains an integrated database of all of the county services (Vaithianathan, et al. 2017).

The researchers built the model analyzing data from all of the CPS referrals from 2008 until 2016 and then included data from other systems including juvenile justice, criminal justice, public benefit programs, census data on neighborhood poverty and behavioral and mental health programs (Vaithianathan, et al., 2017). The result was a model based on 800 different variables from the merged datasets that produced an individual prediction that a child would be either re-referred to CPS within two years or placed into care by CPS within two years (Vaithianathan, et al., 2017). The researchers stated that knowing the possible outcome if a child would be at risk in the future would be the best way to support initial removal decision making (Vaithianathan, et al., 2017).

The AFST process begins when CPS child abuse hotline workers receive a traditional intake and then make a determination about if CPS should open a child abuse investigation case (Chouldechova, et al., 2018). After making that determination, the intake worker then is shown a risk score for that child on a 1 to 20 scale with higher numbers representing higher risk. Scores above 18 are considered a mandatory screen in and intake supervisors are required to override a mandatory score, which happens about a quarter of the time on average, but which differs significantly by each supervisor (Chouldechova, et al., 2018).

In Goldhaber-Fiebert and Prince‘s (2019) impact evaluation of the AFST, they reported that it had a small effect on increasing the number of cases that were screened in for investigation and increased the level of accuracy for both Black and white children, but actually decreased slightly the level of accuracy for cases that were screened out. As a estimate as to how many children might be effected by the model, the researchers guessed that on average each month 24 more children would be accurately screened in by use of the AFST and that 11 would be inaccurately screened out (Goldhaber-Fiebert & Prince, 2019).

How the AFST deals with bias built into the dataset and its effects on minority populations is an ongoing and heated dispute. While acknowledging the inherent racial and human bias in child welfare data, Field et al (2023) looked at recent data from the model and conducted a difference within difference analysis to determine if the AFST reduced racial differences in intake outcomes and in case outcomes. They found that the AFST reduced racial differences in screening outcomes by 2.5 percent and that this effect was most profound for older youth (Field, et al. 2023).

Vaithianathan et al., (2020) tested the AFST on children admitted into the hospital for injuries and found that the model was accurate at predicting injuries associated with abuse. They wrote that the 5 percent of children identified by the model as the highest risk had significantly higher chances of reported injury than those the model identified in the lower 50 percent of risk (Vaithianathan et al., 2020).

However, other researchers have disputed the claim that the AFST reduced bias. When Cheng et al. (2023) looked at the AFST data and interviewed the CPS intake staff workers, they found that the workers themselves did not particularly understand how the AFST model worked, but that they were making conscious choices to screen out more referrals involving black children that the model identified as mandatory. They noted that the AFST model on its own would have screened in 37 percent more black children than what was actually screened in (Cheng, at al, 2022).When the AFST model was tested on a New York State dataset, Du et al. (2022) found that it would increase the number of foster children brought into the system and that it would over identify Black youth.

The ACLU, which has been investigating the AFST, examined the various updated versions of the AFST model and conducted an independent analysis (Gerchick, et al, 2023). Their principal argument against the tool is that the design choices made by its creators perpetuate biases in the system and that those choices are arbitrary and not objective (Gerchick et al. 2023).

The ACLU went on to note that the AFST has three primary problems. It includes some data such as criminal history that is always included no matter the time frame between the child welfare allegation and the past crime (Gerchick, et al, 2023). It also combines scores for an entire family, so risk is calculated for all the children together even if they are different ages. Finally, they wrote that by using as much data as possible for the model, the AFST includes racially biased information such as criminal and juvenile justice data (Gerchick, et al, 2023).

In an interview with Robin Frank, a Pittsburgh attorney who represents families in child welfare cases, Frank stated that attorneys in Allegheny County generally were unaware of the use of the AFST for several years and that it is only recently because of news coverage that the local bar has become more aware of it (Frank, R. personal communication, March 8, 2023). However, she said that Allegheny CPS refuses to share the AFST scores with attorneys or the court, so that attorneys are unaware of what scores their clients might have or what information was used to inform their individual scores. Frank said that while it is possible that attorneys could use discovery to obtain the records, the costs in time and money to do so would generally not be the best use of an attorney's time who is trying to defend a parent or represent a child, so the costs are a barrier to accessing the score for the clients affected.

Ho and Burke (2022) reported that following Associated Press stories about potential bias problems with the AFST, that Oregon CPS stopped using their predictive model that had been based on the AFST. In a subsequent AP story, Ho and Burke (2023) reported that based on complaints that the AFST might also be inadvertently discriminating against protected groups including people with disabilities, the U.S. Department of Justice has begun an investigation into the AFST. However, models based on the AFST are still being used in

Colorado and in California (Ho & Burke, 2023) and a new model is being deployed in Northhampton County, Pennsylvania (Wandalowski, & Vaithianathan, 2023).

What all this research and background reflects is a state of uncertainty and debate over the use of predictive models in the legal community and how the law will deal with them in different legal contexts. How these models might affect child welfare legal decisions is an open question that this paper hopes to help answer by showing that legal removal and placement decisions in the child welfare context can be influenced by the use of predictive models.

METHODOLOGY

As this study was designed to test how legal decisions could be changed by a predictive model, an anonymous survey was sent to law students, attorneys and judges with child welfare and/or juvenile justice experience in the U.S. Recruitment was completed using multiple avenues and began in August 2022 with emails that were sent to child welfare law firms, attorneys working for the state child protective agencies and to attorneys for Court Appointed Special Advocated Guardians Ad Litem (CASA/GAL). The survey was also posted on specialized child welfare and juvenile attorney Facebook groups that had controlled member access. Finally the survey was distributed during two presentations, a webinar for the National Association of Children's Council on February 15, 2023 and during an in person presentation for the National Council of Juvenile and Family Court Judges in Dallas Texas on March 20, 2023. Recruitment concluded in July 2023 with a total of 251 completed responses.

The survey consisted of seven demographic questions followed by instructions to read the following vignettes and then answer the questions as if they were the judge deciding the case. The broad structure of the survey was that participants read and answered questions about two child removal decisions. They were then given randomized high, medium or low predictive model scores and asked to answer the same questions given the new scores. Lastly participants were asked 16 short questions taken from the Dalgleish Scale to measure their beliefs about family preservation and child safety used in multiple studies regarding foster care decision making ((Dalgleish, 2010; Dettlaff, et al., 2020; Hollinshead, et al., 2021).

Prior to distribution, the vignettes and questions were given to five experienced child welfare attorneys with practices in Virginia, Florida and Texas to make certain that the vignette cases were clear and appropriate and that the language of the questions could be understood by participants in different state jurisdictions that might use slightly different terms when discussing the same legal issue. Following qualitative interviews with each attorney, minor changes were made to make the vignettes and questions more universal across U.S. jurisdictions.

This study followed a within subjects design, so all participants saw the same vignettes in the same order. Randomization was done at the treatment level. Participants were first presented with a vignette about a 10-year-old child named Jonas and a removal decision due to issues of neglect and possible medical neglect. The scenario of this vignette was taken from an actual court case in Texas where there was a serious dispute amongst the parties as to removal and placement.

Participants were then asked to answer on a six-point Likert question if they believed Jonas should be removed and taken into state custody with a range from “definitely not” to “definitely”. This question was followed by a five-point Likert question asking participants to rate their confidence in the previous decision with a range of “not at all confident” to “extremely confident”.

Participants then answered two other similar six-point Likert questions regarding if they would place Jonas in a foster home or leave Jonas in his mother’s home. Each of these questions was followed by a five-point Likert questions rating their confidence in each decision.

Following their answers, participants were randomly exposed to one of three treatments, a low, medium or high predictive risk model score. The words “low”, “medium” and “high” were included in the vignette to eliminate any ambiguity about what the numbers themselves meant. The vignette also includes a statement about the accuracy of the algorithm to increase participant confidence.

After their exposure to the treatment, participants were then given the same Likert questions and their corresponding Likert confidence question with the inclusion of the phrase “Given the new predictive risk modeling score . . . “. Questions were given in the same order as before.

Following those questions, participants were given another vignette, this one about a 5-year-old boy named Carlos and possible neglect issues. This vignette was also inspired by a true Texas foster care case where parties disputed the issue of removal and placement.

Participants then followed the same pattern as the first vignette, answering questions and relaying their confidence, then answering the questions again after being exposed to a random, high, medium or low predictive risk model prompt. All told, participants answered 12 questions per vignette, six pretreatment and six post treatment.

Of the 251 completed responses, 208 or almost 83 percent were female, 40 or almost 16 percent were male and two participants chose non binary and one chose not to answer. The overall national percentage of female attorneys in 2022 was 38 percent (American Bar Association, 2022). The uneven distribution in this sample population most likely comes from the specific type of attorney that practices child welfare and juvenile law, which leans heavily on nonprofit, public service and legal aid law firms. The Legal Services Corporation (2021), which is the largest funding organization for legal aid law firms across the United States does track the gender of the lawyers in the legal aid firms it funds and women were the majority of attorneys across the their 899 legal aid offices.

The participant's reported ethnicity is listed below. These figures are similar to the overall national demographics reported by the ABA with white lawyers at 81 percent, Black lawyers 4.7 percent and Hispanic at 5.8 percent(American Bar Association, 2022). the majority of individuals within this sample are White or Caucasian, constituting 77.68% of the total population, followed by Black or African American (7.30%), Hispanic or Latino (6.01%), and Asian or Pacific Islander (3.43%). Additionally, there are smaller percentages of Native American or Alaskan Native (2.58%) and Multi-racial or Biracial (2.15%).

Out of the sample nearly 43 percent had worked as an attorney for 18 or more years, with nearly 30 percent having at least a decade of experience. Private attorneys represent the largest percentage of individuals at 74 responses, followed by 53 attorneys representing CASA or GAL programs and 45 nonprofit legal firm attorneys, and 19 CPS attorneys. In addition, 10 judges, 13 law students and 2 law professors completed the study. Finally, 37 attorneys selected the category *Other Attorney.* Though it is unclear exactly what role those attorneys have in the child welfare system, the survey was also distributed to groups that include attorneys with a primary juvenile justice orientation, so it is likely those responses come from those groups. In any follow up study, juvenile attorneys should be included as their own group. Given that the survey was distributed only to attorneys, judges and law students and there was no financial incentive to take or complete it, it seems unlikely that non-lawyers would have completed the survey. Additionally, time recording data within the survey indicates that these responses match times when other attorney responses were recorded.

Though the survey was sent to groups with nationwide representation and responses were recorded for all but 13 states, a majority of the responses came from Florida with 41, Texas with 40, Arizona with 15, California 15, Colorado 15, Pennsylvania 12 and Georgia 11. I was able to draw from a more robust personal network and thus recruit more attorneys in both Texas and Florida. Since Texas and Florida are the 2nd and 3rd most populous states, the weight of responses from those two states does not seem problematic.

RESULTS

The analysis of the data was done with structural equation modeling (SEM) using Stata to answer the underlying question of if the inclusion of predictive model would affect the legal decision of lawyers and judges regarding initial child removal and placement. (Appendix 1 for regression tables)

The three primary questions following the Jonas vignette were;

1. How likely are you to order that Jonas be taken into the custody of the state protection agency as a ward of the state?

2. How likely are you to order that Jonas be removed and placed in a therapeutic foster home?

3. How likely are you to order that Jonas remain in his mother's home?

The three primary questions following the high, medium or low treatment were;

1. Given the new predictive risk modeling score, how likely are you to order that Jonas be taken into the custody of the state protection agency as a ward of the state?

2. Given the new predictive risk modeling score, how likely are you to order that Jonas be removed and placed in a therapeutic foster home?

3. Given the new predictive risk modeling score, how likely are you to order that Jonas remain in his mother's home?

Jonas Vignette

For each set of Jonas questions, the best predictor of the second response is the what participants selected on their initial response. The coefficient between the two questions are 0.801 for ward of state, 0.76 for foster placement and 0.88 for placement with the mother. This indicates that though the treatments do affect attorney decision making, their initial decision still is the primary predictor.

The total means for each question are as follows; 1st Jonas Ward mean 2.88; 2nd Jonas Ward mean 2.7, 1st Jonas Foster mean 2.77, 2nd Jonas Foster mean 2.68; 1st Jonas Mom mean 3.67,and 2nd Jonas Mom mean 3.70. Beginning with the Jonas questions using the low treatment as the baseline reference category, the results of the model are shown in Table 1.

Table 1 Jonas Treatment Effects

Scenario	Treatment Level	Coefficient (Effect Size)	Statistical Significance	Interpretation
Jonas Ward	Medium	0.365	$p < 0.01$	Increase in willingness to place Jonas under state agency
Jonas Ward	High	0.631	$p < 0.001$	Increase in willingness to place Jonas under state agency
Jonas Ward	Low (vs High)	-0.6108	$p < 0.001$	Less likely to order Jonas into care by state agency compared to high treatment
Jonas Foster Home	Medium	0.17	Not Significant	No significant change in decision to place Jonas in foster home
Jonas Foster Home	High	0.745	$p < 0.001$	More likely to place Jonas into a foster home
Jonas Foster Home	Low (vs Medium)	0.183	Not Significant	Not Significant

Jonas Mom	Medium	-0.401	$p < 0.01$	Less likely to place Jonas back with his mother
Jonas Mom	High	-0.916	$p < 0.01$	Significantly less likely to allow Jonas to stay with his mother
Jonas Mom	Low (vs High)	-0.316	$p < 0.002$	More likely to leave Jonas in his mother's home compared to high treatment

For every Jonas question, high treatments had an effect on legal decision making, meaning lawyers were more likely to support Jonas coming into foster care, being placed in a foster home and not being placed with his mother. Medium treatments also made lawyers more likely to support Jonas being placed into care and less likely to approve placement with the mother. Lawyers who saw the low treatments were less likely to approve of Jonas coming into care and were more likely to approve of placement with his mother.

Carlos Vignette

Similar to the Jonas results, the initial selection by attorneys was the best predictor of what their second response would be. The coefficients for each pair of questions are 0.895 for Carlos being a ward of the state, 0.896 for him being placed into foster care, and 0.848 for being placed with his mother.

The Carlos vignette, had essentially the same questions and structure as the Jonas questions, so the analysis is the same. The total means for each of the questions are; 1st Carlos Ward mean 2.58, 2nd Carlos Ward mean 2.63; 1st Carlos Foster mean 2.37; 2nd Carlos Foster mean 2.52; 1st Carlos Mom mean 3.91; 2nd Carlos Mom mean 3.88. The Carlos question using the low treatment as the baseline reference category results of the model are shown in Table 2.

Table 2 Carlos Treatment Effects

Scenario	Treatment Level	Coefficient (Effect Size)	Statistical Significance	Interpretation
Carlos Ward	Medium	0.383	$p < 0.01$	More likely to support taking Carlos into care of the state agency
Carlos Ward	High	0.808	$p < 0.01$	More likely to support Carlos being taken into care
Carlos Ward	Low (vs High)	-0.794	$p < 0.01$	Less likely to approve of putting Carlos into CPS custody
Carlos Foster	Medium	0.422	$p < 0.01$	More likely to recommend Carlos be placed in a foster home
Carlos Foster	High	0.863	$p < 0.01$	More likely to place Carlos into a foster home
Carlos Foster	Low (vs High)	-0.781	$p < 0.01$	Less likely to place Carlos in the foster home
Carlos Mom	Medium	-0.413	$p < 0.01$	Less likely to place Carlos with his mother
Carlos Mom	High	-0.804	$p < 0.01$	Even less likely to place Carlos with his mother
Carlos Mom	Low (vs High)	0.776	$p < 0.01$	More likely to place Carlos with his mother compared to high treatment

The high treatments for Carlos also had an effect on legal decision making, meaning lawyers were more likely to support Carlos coming into foster care, being placed in a foster home and not being placed with his mother. Medium treatments also made lawyers more likely to support Jonas being placed into care, being placed into a foster home and less likely to approve placement with the mother. Lawyers who saw the low treatments were less likely to approve of Carlos coming into care, being placed in a foster home and were more likely to approve of placement with his mother.

The effects of the treatments were not uniform across the questions and varied between Jonas and Carlos scenarios. The Carlos vignette shows more consistent significant effects across all questions and treatments. The Jonas vignette shows mixed results, with some non-significant effects in the foster placement question, indicating potentially less consistent variation. This indicates that while predictive models can change legal decisions, there effect is not uniform and can depend upon the facts of the case.

In addition to the primary variables, participants were asked demographic questions about their age, race, gender and experience in years. Overall these demographic factors had mixed significant effects on of the questions. Attorneys with the most experience were less likely to place Jonas into care as a ward ($p < 0.05$) . Older attorneys were more likely to leave Jonas with his mom ($p < 0.05$) . Similarly, for the Carlos ward question, the oldest attorneys were less likely to favor placing him into care as a ward ($p < 0.05$) . Taken together, there is a slight indication to suggest that older attorneys might be more skeptical of taking children into care in the context of the vignettes, but the findings are not significant across all the questions, so this interpretation should be taken with caution.

The other somewhat consistent demographic finding is that Black or African American attorneys were more likely to favor both Jonas staying with his mother ($p < 0.05$). They were also more likely to disapprove of Carlos being taken into care as a ward ($p < 0.05$), and being placed into a foster home ($p < 0.05$). Though again, there are not significant finding for all questions and given the small number of Black or African American participants, these should also be interpreted with caution.

Dalgleish Scores

The last explanatory variable tested with the participant's Dalgleish number, representing their internal beliefs about child safety and family preservation.

For the Jonas ward question, the positive coefficient with a p-value of ($p < 0.01$)for Dalgleish number, implies that for each unit increase in the score, there's a slight yet significant increase in the likelihood of recommending Jonas to become a ward of the state. For the Jonas mom placement question, a negative coefficient and p-value ($p < 0.01$) suggests that attorneys with high Dalgleish scores are less likely to favor a placement with the mother.

As for the Carlos vignette, the Carlos ward of the state question had a positive coefficient but was not significant. The Carlos foster care question has a positive coefficient and a p value of ($p < 0.05$), suggesting that as Dalgleish score increases, attorney decisions to place in a foster home is also expected to slightly increase. Finally for the Carlos Mom question with a negative coefficient and p-value ($p < 0.05$) the Dalgleish score has a significant negative effect on decision to place Carlos back with his mom.

All totaled, the Dalgleish score seemed to have only slight impacts, but the direction of the coefficients were similar for both sets of questions indicating that core beliefs about the purpose of foster care do have some effects on decisions in the context of these vignettes.

Confidence Questions

Turning now to the confidence questions, the means for each Jonas question are as follows, 1st Jonas Ward Confidence mean 3.541, 2nd Jonas Ward Confidence mean 3.632; 1stJonas Foster Confidence mean 3.545; 2nd Jonas Foster Confidence mean 3.554; 1st Jonas Mom Confidence mean 3.462; 2nd Jonas Mom Confidence mean 3.549.

The means for each Carlos question are as follows; 1st Carlos Ward Confidence mean 3.648; 2nd Carlos Ward Confidence mean 3.717; 1st Carlos Foster Confidence mean 3.677; 2nd Carlos Foster Confidence mean, 3.629; 1st Carlos Mom Confidence mean 3.593; 2nd Carlos Mom Confidence mean 3.62.

As can be seen by both the central tendency numbers, there is much less change in the confidence questions than in the primary questions. When the confidence questions were run through the same models as the primary questions, none of them produced any significant results. Neither the treatment effects, the demographic questions, nor the Dalgliesh scores were significant for any of the Jonas confidence questions. The Carlos vignette did show some significant effects. For both the ward of the state and the foster placement questions, the medium treatment showed significant effects and the for the mom placement question a higher Dalgleish score was associated with decreased confidence. However, given the lack of significance in any of the Jonas questions and that somewhat strangely, only the medium treatment showed any significance, these results should be interpreted with caution.

DISCUSSION

From the models above, it seems that at least in this context, attorney's legal decisions about child welfare removal and foster care or home placement can be influenced by the inclusion of a predictive model risk score. This is consistent with Engel and Grgić-Hlača's (2021) findings that predictive model scores in a legal context could change people's opinions. However, this is the first time that lawyers and judges have ever been directly tested in a child welfare context.

Though it should not be surprising that a predictive model score can influence legal decision making, the results should be sobering. The decisions regarding removal and placement are often the first ones to be reviewed by lawyers and judges at the beginning of a foster care case and represent essentially the core issue of the legal dispute. The role of judges and attorneys is to make legal decisions and legal recommendations in line with the law, applying the facts of a case to the applicable statutory and case law. The predictive child risk models add new facts to a case that can influence legal conclusions.

The results from the vignette study show small but significant differences between pre and post treatment decisions. This implies that for difficult cases, like the ones presented in the vignettes, a predictive model can sometimes swing removal and placement decisions. While this result might be expected when attorneys saw high and low risk scores, the ability of even medium scores (10) to push legal decisions towards removal and away from placement with the mother, implies that even the quantification of a moderate risk tends to raise concerns about child safety and lead to more conservative decisions. This is however in keeping with the results from Fitzpatrick and Wildman (2021) who found that predictive scores above the median did increase the likelihood that Colorado CPS would begin an investigation. In their study, the mean risk score for a child was just above 8 (Fitzpatrick & Wildman, 2021).

Rachlinki et al. (2015) demonstrated in both criminal and civil cases that judges are subject to the anchoring effect both in damages they award and the criminal sentences they give out. They wrote that judges make unreliable decisions and that "anchoring thus undermines the rule of law by introducing an element of arbitrariness into judicial decisions." (Rachlinki et al., 2015, p. 737). Fraiden (2013) wrote that child welfare judges in particular hear the same case potentially dozens of times and might be subject to a bolstering bias where they weigh their original opinions higher than new evidence.

Taken together, the potential for a predictive model to both anchor and then subsequently affect future decisions raises serious issues for their use in the courtroom. A judge or a CPS attorney responsible for prosecuting the agency's case could be influenced even by a moderate score to order or recommend removal and foster placement and then this score could potentially affect reunification decisions at a future date. The long term effect of predictive models on child welfare decisions is a topic that future studies in the field should consider, but is unknown at the moment.

As discussed in the literature section, these models are often built using historical data that in itself might be inaccurate, biased or not up to date, so a family might get a moderate score with some ease. One of the ACLU criticism of the AFST is that the variable weights used within the model are not published and understood, so how a score is developed is a mystery to those that see it (Gerchick, et al, 2023). In this study, attorneys were presented as part of the vignette with a 90 percent accuracy score for the model, however none of the models discussed so far have an accuracy that high. For instance, the AFST reports an accuracy between 70 and 80 percent (Allegheny County Department of Human Services, 2018). How attorneys might respond to a predictive model with a lower reported accuracy remains another open question for further research.

One general ethics recommendation for the implementation of child welfare predictive models often is the inclusion of stakeholders in the development (Saxena, 2020; Casey Family Programs, 2018; Rahman & Keseru, 2021). In the early creation days of the AFST, stakeholder meetings were held that included court staff and the ACLU, but there is no record of how attorneys responded to the tool (Allegheny County Department of Human Services, 2018). How attorneys perceive these tools and their ethical duties using them remains another open question for further research.

The dearth of attorneys is especially concerning given that this research suggests that these models have the ability to change the outcomes of legal decision making. The literature focus has thus far been limited to discussions of caseworkers and CPS agencies, but all of their decisions are ultimately litigated by attorneys and approved by judges. In the U.S. system, judges are the final decision makers in child welfare. There are no CPS decisions that cannot be approved or overruled by a judge.

In the interview with Pittsburgh child welfare attorney, Robin Frank, she stated that courts and attorneys had been discouraged from inquiring about the AFST number (Frank, R. personal communication, March 8, 2023). This is part of the design of the tool itself (Allegheny County Department of Human Services, 2018). Mills (2019) argued that this lack of transparency was actually a core ethical component of the AFST in that protected judges from being influenced by the score.

However, legal decisions come with legal and ethical duties imposed by the law and shielding judges and attorneys is contrary to both. Lawyers actually have an ethical duty to be involved. ABA Model Rule 1.1 describes that lawyers have a duty of competence and in their commentary the American Bar Association (2019) wrote that this duty includes understanding technology that affects the practice of law and their clients.

ABA Model Rule 8.4 goes further and prohibits an attorney from discriminating against a protected class of persons, which implies that attorneys need to actively understand how models that their clients are using or are being used against their clients might be biased and discriminatory (American Bar Association, 2019). Judicial codes also prohibit judges from being themselves biased against protected categories of people and allowing bias from the attorneys practicing in their court (Texas Code of Judicial Conduct, 2019; Florida Code of Judicial Conduct, 2023).

Finally, attorneys have a duty to under ABA Model Rule 1.3 to diligently represent their clients which included the often used idea of zealous advocacy (American Bar Association, 2023). Zealous representation would at minimum require basic discovery, which is the formal request to the other party for information relevant to the litigation, usually in the form of documents, or written interrogatories, but it can also include sworn deposition testimony. Rule 26 of the Federal Rules of Civil Procedure (2000) states that unless the court orders otherwise, parties are entitled to any nonpriveleged material.

A similar rule can be found in every state. Without delving too deeply into the legal weeds, the duty to represent and the obligations to conduct discovery together mean that attorneys have to ask about any predictive models being used and there is no obvious legal reason why CPS agencies can deny providing them.

To be succinct, this experiment has shown that predictive models can influence child welfare legal decisions. Prior research tells us that judges are subject to the same anchors and biases as the rest of humanity. This concern led model designers to exclude the legal community from much of the knowledge related to their regular use and even though attorneys and judges use them in other legal contexts. The legal community however has an ethical duty to represent their clients, hear cases impartially and be aware of possible biases that technology brings, which ultimately will entail being exposed to predictive model scores.

Though this paper has focused greatly on the AFST, this is just the most timely example as the push to make better decisions of behalf of children will lead to more child welfare predictive model systems being developed around the world. Stakeholder input needs to include the legal community at all stages and it is imperative that the legal community have a greater say in the development and use of predictive models that impact the lives of the clients they represent and the case before the court.

It is also important that the child welfare legal community itself be trained on the uses, potential and biases of predictive models. Arguably the lack of legal attention until now led to the type of circumstances that prompted a Department of Justice investigation. A person would be ill advised to write a contract without consulting a lawyer, certainly deploying a predictive model that can bring the weight of CPS authority upon a family should merit a bit of legal advice.

This present study is limited by the sample size and the amount of judicial participants, so if judges have a different response than attorneys is still unknown. A larger sample of attorneys would also allow subsequent research to determine if the role an attorney plays in the child welfare system affects their response to a predictive model score. Follow up research that focuses on the legal community's responses where these models have been deployed would also be helpful to learn how the models might be affecting attorney representation.

CONCLUSION

Predictive models in many forms are coming for the child welfare field. Many of them are already here. The machine advice they give can change legal conclusions and the legal community has a duty to ethically engage with and debate this advice. This study has implications not only for policy around predictive model use in the child welfare field, but also for legal training for child welfare attorneys and judges. How child welfare legal community thinks about and will respond to this changing technology and how different models will affect their decisions are open research question that needs to be answered hopefully before more of these predictive models are deployed.

REFERENCES

Ahn, E., Gil, Y., & Putnam-Hornstein, E. (2021). Predicting youth at high risk of aging out of foster care using machine learning methods. Child abuse & neglect, 117, 105059. https://doi.org/10.1016/j.chiabu.2021.105059

Allegheny County Department of Human Services, (2018) Frequently Asked Questions. Retrieved October 11, 2023 https://www.alleghenycountyanalytics.us/wp-content/uploads/2018/10/17-ACDHS-11_AFST_102518.pdf

American Bar Association. (2022). ABA Profile of the Legal Profession. Retrieved October 11, 2023https://www.americanbar.org/content/dam/aba/administrative/news/2022/07/profile-report-2022.pdf

American Bar Association (2019) House of Delegates Resolution. Retrieved October 11, 2023https://www.americanbar.org/content/dam/aba/directories/policy/annual-2019/112-annual-2019.pdf

American Bar Association (2023) Model Rules of Professional Conduct, Rule 1.3 Diligence – Comment. https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/rule_1_3_diligence/comment_on_rule_1_3/

Bogert, E., Schecter, A., & Watson, R.T. (2021). Humans rely more on algorithms than social influence as a task becomes more difficult. Scientific Reports, 11. https://doi.org/10.1038/s41598-021-87480-9

Burton, J., Stein, M., & Jensen, T.B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*. https://doi.org/10.1002/bdm.2155

Casey Family Programs. (2018). Considerations for Implementing Predictive Analytics in Child Welfare. Retrieved October 11, 2023 https://www.casey.org/media/Considerations-for-Applying-Predictive-Analytics-in-Child-Welfare.pdf

Cheng, Hao-Fei & Stapleton, Logan & Kawakami, Anna & Sivaraman, Venkatesh & Cheng, Yanghuidi & Qing, Diana & Perer, Adam & Holstein, Kenneth & Wu, Zhiwei & Zhu, Haiyi. (2022). How Child Welfare Workers Reduce Racial Disparities in Algorithmic Decisions. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22). Association for Computing Machinery 1-22. https://doi.org/10.1145/3491102.3501831

Chouldechova, A., Benavides-Prado, D., Fialko, O. & Vaithianathan, R.. (2018). A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, in *Proceedings of Machine Learning Research* 81:134-148 Available from https://proceedings.mlr.press/v81/chouldechova18a.html.

Chor, K. H. B., Luo, Z., Dworsky, A., Raman, R., Courtney, M. E., & Epstein, R. A. (2022). Development and validation of a predictive risk model for runaway among youth in child welfare. *Children and Youth Services Review*, *143*, 106689. https://doi.org/https://doi.org/10.1016/j.childyouth.2022.106689

Collins, D. (2018) 'Predictive policing': Big-city departments face lawsuits, The Associated Press, Retrieved October 11, 2023 https://apnews.com/article/b11e4bca11e548d3af7a63f24e348c6f

Dalgleish, L. (2010). Balance of work focus in child welfare: Work practice and values scales for child protection. Washington, DC: American Humane Association.

Dettlaff, A. J., Hollinshead, D., Baumann, D. J., & Fluke, J. D. (2020). 81Instrumentation to Understand the Child Protective Services Decision-Making Processes. In J. D. Fluke, M. L. López, R. Benbenishty, E. J. Knorth, & D. J. Baumann (Eds.), *Decision-Making and Judgment in Child Welfare and Protection: Theory, Research, and Practice* (pp. 0). Oxford University Press. https://doi.org/10.1093/oso/9780190059538.003.0004

Drake, B., Jonson-Reid, M., Ocampo, M. G., Morrison, M., & Dvalishvili, D. D. (2020). A Practical Framework for Considering the Use of Predictive Risk Modeling in Child Welfare. The Annals of the American Academy of Political and Social Science, 692(1), 162–181. https://doi.org/10.1177/0002716220978200

Du, Y., Ionescu, S., Sage, M., & Joseph, K. (2022). A Data-Driven Simulation of the New York State Foster Care System Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea. https://doi.org/10.1145/3531146.3533165

Elgin, D. J. (2018). Utilizing predictive modeling to enhance policy and practice through improved identification of at- risk clients: Predicting permanency for foster children. *Children and Youth Services Review*, 91, 156–167. https://doi.org/10.1016/j.childyouth.2018.05.030

Engel, C., & Grgic-Hlaca, N. (2021). Machine Advice with a Warning about Machine Limitations: Experimentally Testing the Solution Mandated by the Wisconsin Supreme Court. *Journal of Legal Analysis*. https://doi.org/10.1093/jla/laab001

Federal Rules of Civil Procedure (2000) Rule 26. Duty to Disclose; General Provisions Governing Discovery. Retrieved October 11, 2023. https://www.law.cornell.edu/rules/frcp/rule_26

Field, A., Coston, A., Gandhi, N., Chouldechova, A., Putnam-Hornstein, E., Steier, D., & Tsvetkov, Y. (2023). Examining risks of racial biases in NLP tools for child protective services. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. https://doi.org/10.1145/3593013.3594094

Fitzpatrick, M. & Wildman. C., (2021) Final Report on Douglas County Decision Aid (DCDA) Predictive Risk Modeling Randomized Control Trial Experiment. Retrieved October 11, 2023. https://www.muckrock.com/foi/colorado-127/child-welfare-predictive-analytics-models-122003/

Flores, A., Bechtel, K., & Lowenkamp, C. (2016). False positives, false negatives, and false analyses: Arejoinder to machine bias: There's software used across the country to predict future criminals. Andit's biased against blacks. Federal Probation, 80(2), 38.

(PDF) Algorithmic fairness and resentment. Retrieved October 11, 2023 https://www.uscourts.gov/federal-probation-journal/2016/09/false-positives-false-negatives-and-false-analyses-rejoinder

Florida Supreme Court (2023) Code of Judicial Conduct for the State of Florida. Retrieved October 11, 2023. https://supremecourt.flcourts.gov/Opinions/Judicial-Ethics-Advisory-Committee/Code-of-Judicial-Conduct/Canon-3

Fraidin, M.I. (2013). Decision-Making in Dependency Court: Heuristics, Cognitive Biases, and Accountability. https://ssrn.com/abstract=2358602

Gerchick, M., Jegede, T., Shah, T., Gutierrez, A., Beiers, S., Shemtov, N., Xu, K., Samant, A., & Horowitz, A. (2023). The Devil is in the Details: Interrogating Values Embedded in the Allegheny Family Screening Tool. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. https://doi.org/10.1145/3593013.3594081

Gillingham P. (2016). Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the 'Black Box' of Machine Learning. *British journal of social work*, *46*(4), 1044–1058. https://doi.org/10.1093/bjsw/bcv031

Goldhaber-Fiebert, J. & Prince, L. (2019) Impact evaluation of a predictive risk modeling tool for Allegheny county's child welfare office. Pittsburgh: Allegheny County. Retrieved October 11, 2023. https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/Impact-Evaluation-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL-6.pdf

Grgić-Hlača, N., Castelluccia, C., & Gummadi, K. P. (2022). Taking Advice from (Dis)Similar Machines: The Impact of Human-Machine Similarity on Machine-Assisted Decision-Making. Proceedings of the AAAI Conference on Human Computation and Crowdsourcing, 10(1), 74-88. https://doi.org/10.1609/hcomp.v10i1.21989

Hall, S. F., Sage, M., Scott, C. F., & Joseph, K. (2023). A Systematic Review of Sophisticated Predictive and Prescriptive Analytics in Child Welfare: Accuracy, Equity, and

Bias. *Child and Adolescent Social Work Journal*. https://doi.org/10.1007/s10560-023-00931-2

Hamilton, M. (2018-2019). Debating Algorithmic Fairness. UC Davis Law Review Online, 52, 261-296. Retrieved October 11, 2023. https://lawreview.law.ucdavis.edu/online/vol52/52-online-Hamilton.pdf

Ho, S., & Burke, G. (2023). Here's how an AI tool may flag parents with disabilities. The Associated Press. Retrieved October 11, 2023 https://apnews.com/article/child-protective-services-algorithms-artificial-intelligence-disability-02469a9ad3ed3e9a31ddae68838bc76e

Ho, S., & Burke, G. (2022). Oregon Dropping AI Tool Used in Child Abuse Cases. The Associated Press. Retrieved October 11, 2023. https://pulitzercenter.org/stories/oregon-dropping-ai-tool-used-child-abuse-cases

Hollinshead, D., Currie, D., Kroll, K., Feldman, S., Monahan-Price, K., & Fluke, J. (2021) Associations Between Case, Staff, and Agency Characteristics and the Decision to Place a Child in Out-of-Home Care. Int. Journal on Child Malt. 4, 325–347. https://doi.org/10.1007/s42448-021-00083-9

Keddell, E. (2014). The ethics of predictive risk modelling in the Aotearoa/New Zealand child welfare context: Child abuse prevention or neo-liberal tool? Critical Social Policy, 34(1), 76-96. https://doi.org/10.1177/0261018314543224

Keddell, E. (2019) Algorithmic Justice in Child Protection: Statistical Fairness, Social Justice and the Implications for Practice. Soc. Sci. 8, 281. https://doi.org/10.3390/socsci8100281

Kirk, S. (2015) Children 'not lab -rats' - Anne Tolley intervenes in child abuse, Stuff, Retrieved October 11, 2023. https://www.stuff.co.nz/national/health/70647353/children-not-lab-rats---anne-tolley-intervenes-in-child-abuse-experiment

Lakkaraju, H., Kleinberg, J., & Leskovec, J. (2019). Rethinking "Risk" in Algorithmic Decision-Making. Proceedings of the Conference on Fairness, Accountability, and Transparency, 2019, 89-98. https://doi.org/10.1145/3544548.3581308

Legal Service Corporation (2021) By the Numbers: The Data Underlying Legal Aid Programs. Retrieved October 11, 2023. https://lsc-live.app.box.com/s/e0ejxcp3lp4msvp77mvrzy9w0z354tf7

Leslie, D., Holmes, L., Hitrova, C. & Ott, E. (2020). Ethics review of machine learning in children's social care. [Report] London, UK: What Works for Children's Social Care., Available at SSRN: https://ssrn.com/abstract=3544019

Mills, U. (2019). 'Less Bad' Bias: An analysis of the Allegheny Family Screening Tool. *MPS Seminar One*, Retrieved October 11, 2023. https://medium.com/mps-seminar-one/less-bad-bias-an-analysis-of-the-allegheny-family-screening-tool-ef6ffa8a56fb

Moore, T.D., McDonald, T.P., & Cronbaugh-Auld, K. (2016). Assessing Risk of Placement Instability to Aid Foster Care Placement Decision Making. Journal of Public Child Welfare, 10, 117 - 131. https://doi.org/10.1080/15548732.2016.1140697

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science (New York, N.Y.)*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Purdy, J., & Glass, B.M. (2020). The Pursuit of Algorithmic Fairness: On "Correcting" Algorithmic Unfairness in a Child Welfare Reunification Success Classifier. https://doi.org/10.1016/j.childyouth.2022.106777

Raghavan, M., Barocas, S., Kleinberg, J.M., & Levy, K.E. (2019). Mitigating Bias in Algorithmic Employment Screening: Evaluating Claims and Practices. https://dx.doi.org/10.2139/ssrn.3408010

Rahman, Z., & Keseru, J. (2021). Predictive analytics for children: An assessment of ethical considerations, risks, and benefits1. UNICEF Office of Research - Innocent Working Paper WP 2021-08. https://www.unicef-irc.org/publications/pdf/WP_2021_08.pdf

Raji, I.D., & Fried, G. (2021). About Face: A Survey of Facial Recognition Evaluation. *ArXiv, abs/2102.00813*. https://doi.org/10.48550/arXiv.2102.00813

Rachlinski, J.J., Wistrich, A.J., & Guthrie, C. (2015). Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences. Indiana Law Journal, 90, 6. Retrieved October 11, 2023. https://scholarship.law.cornell.edu/cgi/viewcontent.cgi?article=2552&context=facpub

Samant, A., Horowitz, A., Xu, K., & Beiers, S. (2021). Family surveillance by algorithm: The rapidly spreading tools few have heard of. ACLU: American Civil Liberties Union. https://www.aclu.org/sites/default/files/field_document/aclu_family_surveillance_by_algorithm.pdf

Saxena, D. & Badillo-Urquiola, K. & Wisniewski, P. & Guha, S. (2020). A Human-Centered Review of the Algorithms used within the U.S. Child Welfare System. https://doi.org/10.1145/3313831.3376229

Saxena, D., Repaci, C., Sage, M., & Guha, S. (2022). How to Train a (Bad) Algorithmic Caseworker: A Quantitative Deconstruction of Risk Assessments in Child Welfare. CHI Conference on Human Factors in Computing Systems Extended Abstracts. https://doi.org/10.48550/arXiv.2203.05695

Stepura, K. G., Schwab, J., Baumann, D. J., Sowinski, N., & Thorne, S. (2021). Exploration in Predictive Analysis and Machine Learning In: Decision-Making and Judgment in Child Welfare and Protection. In J. D. Fluke, M. López López, R. Benbenishty, E. J. Knorth, & D. J. Baumann (Eds.), Oxford University Press. https://doi.org/10.1093/oso/9780190059538.003.0002

Texas Supreme Court (2019) Texas Court Rules - Code of Judicial Conduct. Retrieved October 11, 2023. https://www.scjc.texas.gov/media/46779/code-of-judicial-conduct.pdf

Vaithianathan, R., Jiang, N., Maloney, T., Nand, P., & Putnam-Hornstein, E. (2017). Developing predictive risk models to support child maltreatment hotline screening decisions: Allegheny County methodology and implementation [PDF]. Auckland: Centre for Social Data Analytics. Retrieved October 11, 2023. https://www.alleghenycountyanalytics.us/wp-content/uploads/2019/05/Methodology-V1-from-16-ACDHS-26_PredictiveRisk_Package_050119_FINAL.pdf

Vaithianathan, R., Putnam-Hornstein, E., Chouldechova, A., Benavides-Prado, D., & Berger, R. (2020). Hospital Injury Encounters of Children Identified by a Predictive Risk Model for Screening Child Maltreatment Referrals: Evidence From the Allegheny Family Screening Tool. JAMA Pediatrics, 174(11), e202770-e202770. https://doi.org/10.1001/jamapediatrics.2020.2770

Vaithianathan, R., Dinh, H., Kalisher, A., Kithulgoda, C., Kulick, E., Mayur, M., . . . Prado, D. B. (2019). Implementing a child welfare decision aide in Douglas County: Methodology report (Tech. Rep.). Auckland, New Zealand: Centre for Social Data Analytics. Retrieved October 11, 2023. https://csda.aut.ac.nz/__data/assets/pdf_file/0009/347715/Douglas-County-Methodology_Final_3_02_2020.pdf

Vaithianathan, R., Cuccaro-Alamin, S., Putnam-Hornstein, E. (2023). Improving Child Welfare Practice Through Predictive Risk Modeling: Lessons from the Field. In: Connell, C.M., Crowley, D.M. (eds) Strengthening Child Safety and Well-Being Through Integrated Data Solutions. Child Maltreatment Solutions Network. Springer, Cham. https://doi.org/10.1007/978-3-031-36608-6_8

Wandalowski, S., Vaithianathan, R. (2023). Northampton County: Pursuing Better Child Welfare Outcomes with a Decision Aid Tool to Support Caseworkers. In: Shenk, C.E. (eds) Innovative Methods in Child Maltreatment Research and Practice. Child Maltreatment Solutions Network. Springer, Cham. https://doi.org/10.1007/978-3-031-33739-0_12

Willcott, A. & Stewart, L. (2021) Cold Case Project, Annual Report of the Work of 2019, Georgia's Child Welfare System, Retrieved October 11, 2023. http://www.gacip.org/cold-case-project/

APPENDIX

VARIABLES	(1) Jonas_2_Ward	(2) /	(3) Jonas_2_Mom	(4) /	(5) Jonas_2_Foster_	(6) /
How likely are you to order that Jonas be taken into the custody of the state	0.749***					
	(0.666 - 0.832)					
Medium	0.365***		-0.401***		0.174	
	(0.131 - 0.599)		(-0.624 - -0.179)		(-0.062 - 0.410)	
High	0.631***		-0.916***		0.745***	
	(0.401 - 0.861)		(-1.139 - -0.694)		(0.512 - 0.977)	
Years Since First Licensed as Attorney = 3-6 years	-0.041		-0.025		0.425	
	(-0.770 - 0.688)		(-0.731 - 0.682)		(-0.316 - 1.165)	
Years Since First Licensed as Attorney = 6-9 years	-0.440		0.516		-0.086	
	(-1.144 - 0.264)		(-0.164 - 1.197)		(-0.802 - 0.630)	
Years Since First Licensed as Attorney = 9-12 years	-0.438		0.047		0.244	
	(-1.156 - 0.280)		(-0.647 - 0.742)		(-0.485 - 0.973)	
Years Since First Licensed as Attorney = 12-15 years	-0.196		0.080		0.320	
	(-0.931 - 0.539)		(-0.629 - 0.790)		(-0.427 - 1.067)	
Years Since First Licensed as Attorney = 15-18 years	-0.601		0.299		-0.029	
	(-1.328 - 0.125)		(-0.403 - 1.002)		(-0.767 - 0.709)	

Years Since First Licensed as Attorney = 18+ years	-0.640*		0.369		-0.075	
	(-1.342 - 0.063)		(-0.309 - 1.048)		(-0.789 - 0.639)	
Years Since First Licensed as Attorney = Not licensed/Law student	-0.487		0.222		0.355	
	(-1.186 - 0.213)		(-0.452 - 0.897)		(-0.352 - 1.062)	
How old are you? = 25-34 years old	0.402		0.261		-0.206	
	(-0.330 - 1.134)		(-0.445 - 0.966)		(-0.951 - 0.540)	
How old are you? = 35-44 years old	0.453		0.338		-0.226	
	(-0.332 - 1.238)		(-0.416 - 1.093)		(-1.025 - 0.572)	
How old are you? =, 45-54 years old	0.633		0.183		-0.191	
	(-0.143 - 1.409)		(-0.565 - 0.932)		(-0.980 - 0.599)	
How old are you? = 55-64 years old	0.685*		-0.021		-0.016	
	(-0.126 - 1.496)		(-0.802 - 0.760)		(-0.840 - 0.809)	
How old are you? = 65+ years old	0.526		0.189		-0.054	
	(-0.330 - 1.382)		(-0.635 - 1.013)		(-0.925 - 0.816)	
Race/Ethnicity -,Black or African American	-0.076		0.561*		0.268	
	(-0.706 - 0.555)		(-0.045 - 1.167)		(-0.371 - 0.906)	
Race/Ethnicity - Hispanic or Latino	0.012		0.445		0.246	
	(-0.615 - 0.639)		(-0.160 - 1.051)		(-0.390 - 0.883)	
Race/Ethnicity - Native American or Alaskan Native	-0.220		0.733*		-0.480	
	(-1.008 - 0.568)		(-0.027 - 1.493)		(-1.280 - 0.321)	
Race/Ethnicity - White or Caucasian	0.219		0.384		0.338	
	(-0.308 - 0.745)		(-0.124 - 0.893)		(-0.197 - 0.873)	
Race/Ethnicity - Multi-racial or Biracial	-0.055		0.750*		0.105	
	(-0.867 - 0.758)		(-0.033 - 1.533)		(-0.721 - 0.931)	
Race/Ethnicity - A race or ethnicity not listed here.	-0.076		0.865		-0.388	
	(-1.224 - 1.072)		(-0.242 - 1.972)		(-1.554 - 0.778)	
Which of the following best describes you? -, Male	0.230		-0.256*		0.193	
	(-0.052 - 0.512)		(-0.527 - 0.014)		(-0.093 - 0.480)	
Which of the following best describes you? - Non-binary	-0.174		0.358		0.096	
	(-1.229 - 0.881)		(-0.662 - 1.378)		(-0.980 - 1.171)	
Which of the following best describes you? - Prefer not to answer	-0.101		0.241		-0.184	
	(-1.515 - 1.313)		(-1.123 - 1.606)		(-1.620 - 1.252)	
Dalsum	0.011***		-0.011***		0.003	
	(0.004 - 0.017)		(-0.017 - -0.005)		(-0.004 - 0.009)	
var(e.Jonas_2_Ward)		0.485***				
		(0.397 - 0.574)				
How likely are you to order that Jonas remain in his mother ’home?			0.765***			
			(0.690 - 0.839)			
var(e.Jonas_2_Mom)				0.452***		
				(0.369 - 0.534)		
How likely are you to order that Jonas be removed and placed in a therapeutic					0.749***	
					(0.670 - 0.828)	
var(e.Jonas_2_Foster_)						0.501***
						(0.409 - 0.592)
Constant	-0.002		0.515		0.070	
	(-1.067 - 1.064)		(-0.531 - 1.561)		(-1.008 - 1.149)	
Observations	231	231	231	231	231	231

VARIABLES	(1) Carlos_2_Ward	(2) /	(3) Carlos_2_Foster	(4) /	(5) Carlos_2_Mom	(6) /
How likely are you to order that Carlos be taken into the custody of the state?	0.871***					
	(0.036)					
Medium	0.383***		0.422***		-0.413***	
	(0.093)		(0.094)		(0.119)	
Large	0.808***		0.863***		-0.804***	
	(0.099)		(0.100)		(0.126)	
Years Since First Licensed as Attorney = 3-6 years	-0.222		0.212		-0.025	
	(0.302)		(0.302)		(0.384)	
Years Since First Licensed as Attorney = 6-9 years	-0.134		0.175		-0.297	
	(0.291)		(0.292)		(0.371)	
Years Since First Licensed as Attorney = 9-12 years	0.138		0.328		-0.402	
	(0.298)		(0.297)		(0.379)	
Years Since First Licensed as Attorney = 12-15 years	-0.125		0.341		-0.236	
	(0.304)		(0.304)		(0.388)	
Years Since First Licensed as Attorney = 15-18 years	-0.230		0.078		-0.003	
	(0.299)		(0.299)		(0.383)	
Years Since First Licensed as Attorney = 18+ years	-0.043		0.133		-0.145	
	(0.290)		(0.291)		(0.370)	
Years Since First Licensed as Attorney = Not licensed/Law student	-0.090		0.307		-0.157	
	(0.287)		(0.288)		(0.367)	
How old are you? =25-34 years old	0.096		0.069		-0.463	
	(0.302)		(0.302)		(0.385)	
How old are you? = 35-44 years old	0.158		0.121		-0.265	

	(0.323)		(0.322)		(0.411)	
How old are you? = 45-54 years old	0.085		0.053		-0.199	
	(0.319)		(0.320)		(0.408)	
How old are you? = 55-64 years old	-0.071		-0.018		-0.320	
	(0.335)		(0.334)		(0.426)	
How old are you? = 65+ years old	0.201		0.005		-0.654	
	(0.350)		(0.351)		(0.447)	
Race/Ethnicity - Black or African American	-0.630**		-0.499*		0.519	
	(0.258)		(0.259)		(0.331)	
Race/Ethnicity - Hispanic or Latino	-0.451*		-0.391		0.768**	
	(0.257)		(0.258)		(0.329)	
Race/Ethnicity - Native American or Alaskan Native	-0.396		-0.496		0.898**	
	(0.324)		(0.326)		(0.415)	
Race/Ethnicity - White or Caucasian	-0.508**		-0.327		0.800***	
	(0.216)		(0.216)		(0.276)	
Race/Ethnicity - Multi-racial or Biracial	-0.614*		-0.133		0.499	
	(0.334)		(0.334)		(0.426)	
Race/Ethnicity - A race or ethnicity not listed here.	-0.991**		-0.739		0.365	
	(0.470)		(0.472)		(0.601)	
Which of the following best describes you? - Male	-0.146		0.119		-0.204	
	(0.113)		(0.113)		(0.144)	
Which of the following best describes you? - Non-binary	0.520		0.046		-0.345	
	(0.437)		(0.439)		(0.558)	
Which of the following best describes you? Prefer not to answer	0.018		-0.087		-0.371	
	(0.583)		(0.585)		(0.746)	
Dalsum	0.004		0.006**		-0.008**	
	(0.003)		(0.003)		(0.004)	
var(e.Carlos_2_Ward)		0.317***				
		(0.029)				
How likely are you to order that Carlos be removed and placed in a foster home?			0.857***			
			(0.039)			
var(e.Carlos_2_Foster)				0.319***		
				(0.030)		
How likely are you to order that Carlos remain in his mother's home?					0.794***	
					(0.040)	
var(e.Carlos_2_Mom)						0.518***
						(0.048)
Constant	0.508		0.145		0.978*	
	(0.433)		(0.434)		(0.580)	
Observations	231	231	231	231	231	231