A Data-Driven Simulation of the New York State Foster Care System

Yuhao Du
yuhaodu@buffalo.edu
University at Buffalo
Buffalo, USA

Stefania Ionescu
ionescu@ifi.uzh.ch
University of Zurich
Zurich, Switzerland

Melanie Sage
msage@buffalo.edu
University at Buffalo
Buffalo, USA
msage@buffalo.edu

Kenneth Joseph
kjoseph@buffalo.edu
University at Buffalo
Buffalo, USA

ABSTRACT

We introduce an analytic pipeline to model and simulate youth trajectories through the New York state foster care system. Our goal in doing so is to forecast how proposed interventions may impact the foster care system's ability to achieve it's stated goals *before these interventions are actually implemented and impact the lives of thousands of youth*. Here, we focus on two specific stated goals of the system: racial equity, and, as codified most recently by the 2018 Family First Prevention Services Act (FFPSA), a focus on keeping all youth out of foster care. We also focus on one specific potential intervention— a predictive model, proposed in prior work and implemented elsewhere in the U.S., which aims to determine whether or not a youth is in need of care. We use our method to explore how the implementation of this predictive model in New York would impact racial equity and the number of youth in care. While our findings, as in any simulation model, ultimately rely on modeling assumptions, we find evidence that the model would not necessarily achieve either goal. Primarily, then, we aim to further promote the use of data-driven simulation to help understand the ramifications of algorithmic interventions in public systems.

CCS CONCEPTS

• **Computing methodologies** → **Model verification and validation**.

KEYWORDS

Simulation, Child Welfare, Social Policy, Racial Equity

ACM Reference Format:
Yuhao Du, Stefania Ionescu, Melanie Sage, and Kenneth Joseph. 2022. A Data-Driven Simulation of the New York State Foster Care System. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), June 21–24, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3531146.3533165

FAccT '22, June 21–24, 2022, Seoul, Republic of Korea

ACM ISBN 978-1-4503-9352-2/22/06…$15.00
https://doi.org/10.1145/3531146.3533165

1 INTRODUCTION

Youth admitted into foster care in the United States are likely to experience a series of investigations and evaluations. This constant surveillance, often combined with frequent and abrupt shifts in living situations, leave a number of lasting socio-emotional scars [27, 48]. These scars are, moreover, not distributed equally. In particular, significant racial biases exist regarding who enters into foster care [5, 22].

These and other issues with the American child welfare system have led to a growing movement to abolish it [11]. Suggested alternatives from these advocates include reallocating funding to community-based, localized initiatives not run by government actors [11]. Others have, in a more traditional vein for the field of Social Work, instead argued for changing, rather than eradicating the system. However, both abolitionists and reformers generally agree on one point: youth are better off with their own families than in the foster care system [6].

The most important recent development aimed at keeping youth out of care is the *Family First Prevention Services Act (FFPSA)* [32]. The FFPSA, signed into federal law in 2018, overhauled funding in the American child welfare system in order to providing financial incentives to states to keep youth out of foster care, or more formally, to 1) reduce unnecessary separation, and 2) provide family-based service for removed youth in order to encourage family reunification. These incentives include the construction of new benchmarks for numbers of youth in care, and additional funding for family-based services aimed specifically at helping the families of youth who are in care to get to a point where the youth can be returned to them.

Efforts to effectively use this funding from the FFPSA to reduce the number of youth in foster care face three formidable challenges. First, the American child welfare system is heavily decentralized, with drastically different policy environments across and often even within states [50]. Changes that are effective at moving towards the goals of the FFPSA in one county or state may therefore be difficult (or even illegal) to enact in others. Second, even in a single jurisdiction, the process of placing a youth into the foster care system is complex. When a potentially maltreated youth is reported

to Child Protective Service (CPS), a series of decisions made by case workers and judges are carried out to decide the best path for the youth. The process is cumbersome and riddled with the potential of both personal and systemic biases [36, 53]. Efforts to discharge youth from foster care can face similar challenges. Finally, as with most social policy settings [37], attempts to make progress on one goal can often have unexpected negative impacts on others. As we highlight here, for example, interventions aimed making the system more racially equitable may end up leading to substantially more youth in foster care, and not with their families.

Help is needed to tackle these challenges. Abdurahman [1] argues that implicit in the FFPSA is the assumption that such help will come from the expansion of data collection the bill enables, which will in turn lead to the construction of (predictive) analytical tools that can serve to ameliorate potential problems. Indeed, the use of such tools is already widespread in child welfare, and is only expected to grow with the FFPSA [42].

The present work takes a different bend on how computation can serve a role in social change [2], arguing that if used carefully, computational models can help us to reason about the best path forward under myriad possible social and policy environments, and various potential interventions. That is, rather than make predictions about youth, we can use computation to help us *understand* the foster care system as a whole, and to *rethink* the potential interventions needed to address new policy goals like those set forth in the FFPSA.

The present work conducts such an analysis using a two-stage pipeline. We first take a *forensic social science* [31] approach to address the following research question: **how do existing practices within the New York State child welfare system fare in light of the new stated goals introduced by FFPSA and the existing stated goal of racial equity?** While our methodology generalizes to other contexts, we focus on New York because of our policy and practical expertise in the state. Forensic social science is a methodology in which computational analyses of observational data are conducted in ways that inform and are informed by relevant social theory. Here, using complete data from the child welfare system in New York from 2000-2017 [34], we conducted a computational analysis on 1) entry rates of youth into the foster care system, 2) patterns in how long youth stay in care, and 3) the rates at which youth are discharged. We also consider how these quantities differ for white versus Black youth.

Our forensic social science analysis informs the second, and focal, part of our computational pipeline, in which we construct a *data-driven system dynamics simulation* [17, 47] to analyze a hypothetical intervention into New York's foster care system. Specifically, we analyze how the introduction of an automated risk assessment tool currently in practice in other parts of the country [8] would, if implemented in New York, help to address the goals of the FFPSA while maintaining the existing goal of improving racial equity. System dynamics models help formalize how particular entities *flow* through a system over time, conditioned on assumptions about the probability of flowing from one point to another. Here, we model the flow of American youth into and out of foster care. In a data-driven system dynamics model, some of these patterns of flow, and associated probabilities, can be informed by data. Those that cannot be informed by available data can then be set based on assumptions and/or theory. These assumptions, in turn, can then be varied to address certain research questions or to perform robustness checks.

Our work makes three primary contributions:

(1) We provide quantitative evidence, supporting earlier work, that show that youth stays in foster care in New York can be empirically separated into two classes, long-term and short-term. We find further that the duration of long-term, but not short-term, stays varies significantly across racial lines for all ages.
(2) Informed by these analyses, we develop a data-driven simulation model of the U.S. foster care system, and parameterize it for the study of New York. As a sign of the model's validity, we show that the model can reliably forecast patterns in the number of youth admitted into New York state's foster care system in 2018, given only data from previous years.
(3) We use our model to study the impacts of a potential algorithmic intervention in New York. Our observations resonate with the concerned pointed out by Samant et al. [42]. Specifically, we find that it is difficult to balance the goals of racial equity and a reduction of youth in care, and that proposed algorithmic interventions—encouraged by the FFPSA in order to achieve it's goals [1]—are not necessarily capable of doing so.

As with any simulation model, our findings rely on modeling assumptions which may be reasonably disagreed with. To facilitate such discussions, we have made our model publicly available.[1] However, as we believe our modeling assumptions to be at least within the realm of possibilities, our results suggest that shifting policy landscapes impacts both the validity and utility of using historical administrative dataset to build machine learning models in public sectors. Instead, we argue that machine learning, and computation writ large, may be better served as a tool to facilitate social theory and social policy, rather than to act as explicitly as a decision-making tool, where it is often inserted into the problematic decision loops which expose risks of amplifying existing problems.

2 RELATED WORK

Computational social scientists have develop a number of ways to formalize and analyze complex sociotechnical systems [17, 28, 45]. Careful formalization and analysis can serve to illuminate and identify paths towards addressing societal issues [13]. The benefits and drawbacks of computational modeling of (and for) social systems are perhaps best summarized by Abebe et al. [2], who note that computing can act as a synecdoche, allowing us to think about problems in new ways, and as a rebuttal, "clarify[ing] the limits of technical interventions" [2, pg. 256]. Our goal in the present work, similar in some respects to the arguments made by Green [20] but with distinct methods, is to use computation as both a synecdoche and as a rebuttal for the blind reapplication of machine learning methods from one context to another.

Here, we define computational analysis broadly, to include both forensic social science and simulation. With respect to the former, forensic social science entails the combined use of machine learning and social theory to advance our understanding of social phenomenon, where (in the forensic social science approach)

[1] https://github.com/yuhaodu/system_dynamic_simulation_FC

machine learning/advanced statistical analyses are applied an "atheoretical [approach to] agnostic search for potential explanations," and social theory is "a focusing device that identifies which constructs are to be selected and formed from the millions of possible [analyses]" [31], (pg. 10). We briefly touch on related work in the context of quantitative studies of American Foster Care System that inform our forensic social science analysis, and the simulations of sociotechnical systems that compliment and/or inform the work presented here.

2.1 Quantitative Study of the American Foster Care System

Quantitative studies of foster care are routinely conducted by and for policy makers, case workers and researchers. Given the rich administrative datasets curated in this context, quantitative studies have been carried out to both understand the different aspects of the system and to provide assistance with decision making.

With respect to quantitative work that assists decision making, we direct the reader to a number of recent literature reviews [42, 44] around algorithms used within the U.S. Child Welfare System. Most notably here, Chouldechova et al. [8] build and implement a machine learning model called AFST in Allegheny County, Pennsylvania that helps case workers decide whether or not to screen in reported cases of child abuse for further analysis. This academic work is complemented by a number of public reports as well (e.g. [18]).

A number of other researchers have used quantitative methods to critique existing practices within the U.S. foster care system (e.g. [35]), to better understand associations with service allocation to youth (e.g. [29, 52]), to analyze racial disparities within the system (e.g. [7, 25, 51]), and to criticize the use of automatic decision tools in foster care system (e.g. [10, 14, 38]). Our work compliments these efforts, both in its use of quantitative methods to explore youth lengths of stay in new ways, and to use quantitative methods to critique existing practices.

2.2 Social Simulation

The present work uses a simulation methodology particularly well-suited for our work: *data-driven simulation*. The term data-driven simulation encapsulates a broad range of computing techniques which use relevant data to make educated predictions about what might happen in a situation for which complete data cannot be obtained. Data-driven simulation has seen increasing use in the FAccT-aligned community [9, 19, 23, 24, 33, 33] As noted above, we use a specific form of data-driven simulation, System Dynamics modeling [47]. Perhaps most relevant to our work, then, Martin et al. [30] argues that instead of focusing on mathematical-based interventions on opaque algorithms and/or models, using community based system dynamics modeling to place the algorithm/model into the social context is a better way to understand the long-term impact of algorithms/model. While we aspire to community-based methods, our current work relies only on our existing knowledge of and experiences as practitioners within child welfare.

Finally, while significantly distinct in focus, it is worth noting that there is other work using simulation in the foster care context. Specifically, Fowler et al. [15] use a system dynamic model to test the impact of scaling up a policy to provide long-term rental subsidies for foster care family, and Goldhaber-Fiebert et al. [19] use decision-analytic model in support of Child Welfare policymakers considering implementing evidence-based interventions.

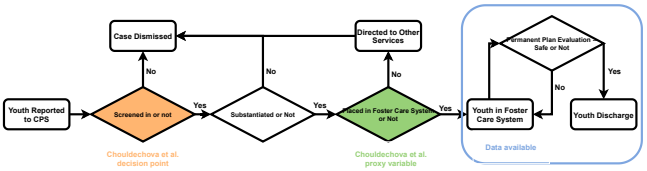


Figure 1: The standard decision-making pipeline for a youth reported to Child Protective Services in the United States. Diamonds represent decision points in the system, squares are states that youth may be in at a given time. The blue box represents the portion of the decision pipeline informed by AFCARS data. The orange and green boxes represent, respectively, the decision point for the algorithmic intervention considered in the present work, and the proxy variable used to train the proposed model.

3 A BRIEF OVERVIEW OF HOW YOUTH ENTER FOSTER CARE IN AMERICA

Our work assumes a slightly simplified process model of youth trajectories through the U.S. foster care system introduced by Sterman [47], and visualized in Figure 1. The first step in the process of placing a youth in foster care is a report to Child Protective Services (*CPS*). These reports are handled by a CPS employee who decides whether or not to *screen in* the call, a decision based on myriad factors, e.g. characteristics of the youth, their family, and the local policy environment [8].

If a call is not screened in, the case is dismissed, and the youth exits the system (although notably, their data may not [1]). If a call is screened in, the case is then taken on for further consideration by a CPS case worker. This phase typically includes a more detailed records review of the youth and their family, and a visit to the location relevant to the call (often, the current living situation of the youth). The CPS case worker then decides whether or not the case is *substantiated*, i.e. whether there is evidence that the concern voiced in the original call is true.

If a case is substantiated, a decision is then made on whether to a) remove the youth from their home, or b) to keep the youth in their home. This decision is typically made by a judge. If the decision is that the youth should be removed from their family, the youth then enters foster care. Once the youth is in foster care, the family is repeatedly re-evaluated for a need to be in the system. More specifically, judges and case workers are expected to return the youth to their family when they are satisfied that it is safe to do so, while also making a backup plan, according to what is referred to as the permanency plan for that youth. When the home is safe again, or alternatively another long-term plan such as guardianship or adoption is available, the youth is then discharged back to a living placement outside of the foster care system.

4 DATA

Our analysis uses federal administrative data from the *Adoption and Foster Care Analysis Reporting System (AFCARS)* [34], a dataset from the *National Data Archive on Child Abuse and Neglect (NDACAN)*. AFCARS data contains a range of information for all foster youth from all 50 U.S. states, the District of Columbia, and Puerto Rico. This information includes demographic data, like race and gender, and administrative data, like the youth's current placement setting. However, critically, such information is only available for youth who *are in foster care*. Other relevant data — for example, the number of youth who are reported but not removed, are not in this dataset. As we discuss further below, we can thus only make assumptions about these youth, which we can then vary to emulate different potential real-world settings.

Our analysis uses a particular sample of youth from the full AFCARS dataset. First, as noted above, we focus our analysis only on data from New York state. Second, our analysis below considers racial disparities in a number of ways. Due to limited data availability leading to imprecise estimates, we focus here only on data for youth who identify (or are identified as) white or Black. This is a significant limitation of the present work that could be alleviated in the future by additional data and more diverse expertise.

5 FORENSIC SOCIAL SCIENCE ANALYSIS

We conduct a forensic social science analysis to analyze the foster care system in New York with respect to the number of youth in care and racial equity within the system. These analyses are useful both in providing a better understanding of how many youth are in care at any time and for how long they remain in foster care, for seeing how these quantities vary for Black versus white youth, and for informing parameters of our data-drive simulation model.

We conduct an analysis of three quantities: how many (Black vs. white) youth are entering into care (the *entry rate*), how long youth stay in care (*length of stay*), and the rate at which youth are discharged (*discharge rate*). Prior work has analyzed entry rates at the state level, finding potential associations with race and age [4, 41]. Other work analyzing youth in Florida also suggest race is a key factor influencing the length of stay [4]. Motivated by this work, we therefore study these quantities split out by youth age in all cases. Moreover, in order to obtain results which are informative for the system dynamics model, we aggregate data at a monthly level.

5.1 Entry Rate

To determine entry rates, we extract the total number of admitted youth $n^t_{r,a}$, where r represents the youth's race (here, Black or white), a represents the age at which they were admitted (rounded down), and t represents the admitted month. For each combination of r and a, we calculate the difference of the number of admitted youth between two consecutive months by $\Delta^t_{r,a} = n^{t+1}_{r,a} - n^t_{r,a}$. After confirming that this difference is stationary for the vast majority of (r, a) combinations[2], we fit the resulting data to normal distributions $\mathcal{N}^{\Delta_{r,a}}$ for each $\Delta_{r,a}$ along t.

[2]Using the Augmented Dickey-Fuller test [12], only one p-value was greater than .05, and only two were greater than .001.

Age	Distribution (Black)	Distribution (white)
1	$\mathcal{N}(0.25, 7.1)$	$\mathcal{N}(-0.05, 6.0)$
5	$\mathcal{N}(0.15, 5.2)$	$\mathcal{N}(0.01, 5.8)$
9	$\mathcal{N}(0.1, 3.8)$	$\mathcal{N}(0.04, 4.1)$
13	$\mathcal{N}(0.09, 5.9)$	$\mathcal{N}(0.025, 4.8)$

Table 1: Fitted normal distributions for the difference between the number of admitted Black or white youth in consecutive months in the AFCARS dataset

Samples of the estimated distributions for changes in entry rates for white and Black youth at various ages are displayed in Table 1. Across the visualized quantities, as well as for all ages not displayed, we find no significant differences in the *rate of change* in the number of Black and white youth entering the system. However, consistent with prior work at a national level [51], we observe that Black youth are over-represented relative to white youth within New York.

5.2 Length of Stay

Figure 2a shows the density of the *logarithm of the length of stay (log-LOS)* for white and Black youth admitted at ages 1, 5, 9, and 13. Other ages show similar patterns, and thus we omit the results here. From these plots, we notice that the distribution of log-LOS appears to be readily modeled by a mixture of Gaussian distributions. In other words, there appear to be distinct *classes* of stay lengths for foster youth, that align with different distributions of length of stay.

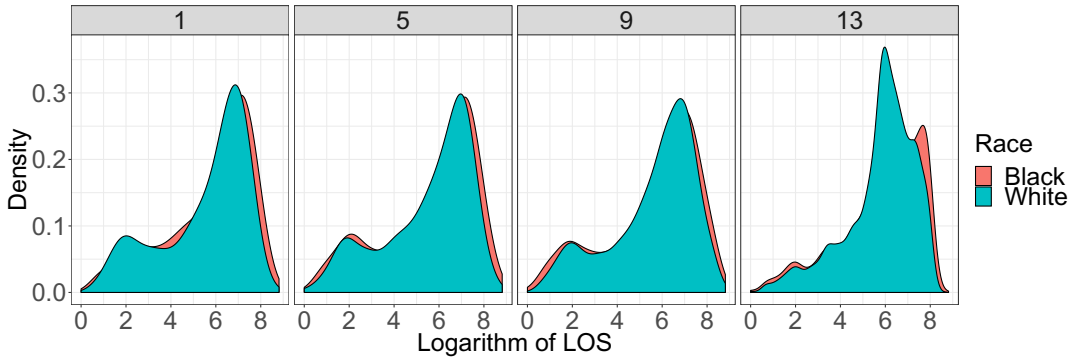
To test this observation statistically, we fit each individual distribution of log-LOS for all combinations of a and r to five different Gaussian mixture models, with the assumed number of distributions N to be either $N = 1, 2, 3, 4, 5$. Confirming our intuitions from Figure 2a, we find that $N = 2$ best fitted all densities for all age and race combinations according to the AIC score.

This Gaussian mixture model separates stays in foster care into two groups- *long-term* and *short-term* stays. Long-term stays have longer log-LOS than short-term stay. For example, among white youth at age 10, the median length of a short-term stay is 10 days, while it is 680 days for a long-term stay. Among Black youth at age 10, the median length of a short-term stay is 9 days, and 804 for a long-term stay.

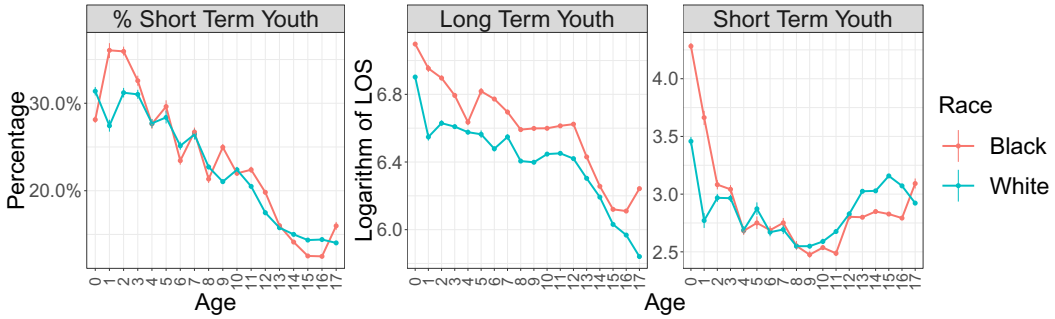
Breaking AFCARS data down further by long-term youth vs short-term youth is revealing. First, as shown in the left-most plot in Figure 2b, we find that between 10-30% of admitted Black and white youth are short-term, with that quantity generally (and predictably [52]) decreasing as youth are older. Second, the middle subplot shows that, consistently across age groups, long-term stays for Black youth are longer than those for white youth. Similarly, the rightmost plot in Figure 2b, shows that that from ages 0-2, short-term stays are also longer for Black youth. There is some evidence that short term stays are slightly longer for older white youth, but this finding is inconsistent at older ages (notably, at age 17).

5.3 Discharge Rate

We formalize discharge rate as *the percentage of youth that are in care after t months*. We again learn this function separately for all combinations of a and r, as well as for long/short-term foster care youth. We then extracted the empirical inverse cumulative



(a) The distributions of log-LOS depending on the youth's race. From left to right, plots correspond to a different age at which the youth was admitted into foster care: 1, 5, 9, and 13. Other ages show similar patterns, and are not displayed



(b) On the X-axis is youth's admitted age. The Y-axis of the leftmost plot represents the percentage of short-term youth and the Y-axis at the right two plots represents mean of log-LOS. Cyan lines shows results for white youth, red for Black youth. Data are for youth in New York state in the AFCARS dataset. Confidence interval are 95% bootstrapped CIs.

Figure 2: Split of youth according to their LOS

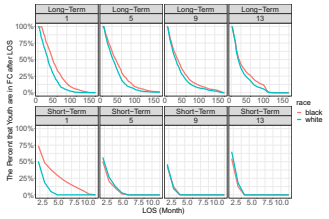


Figure 3: On the X-axis is youth length-of-stay. The Y-axis represents the probability that a given youth will still be in foster care after spending the given number of months in foster care. The first row shows the results for long-term youth while second row shows the results for short-term youth. Cyan lines shows results for white youth, red for Black youth. Data are for youth in New York state in the AFCARS dataset

distribution of LOS for long/short-term stays for youth of different demographics. Finally, we learn a non-parametric discharge function using a linear interpolation to fit the empirical inverse cumulative distribution. This *discharge function* $\mathcal{D}^{l}_{r,a}(d_t)$ gives us the probability of l-term youth whose race is r and admitted age is a will still be in foster care after d_t months.

A representative portion of the results from this analysis are presented in Figure 3, which shows the expected length of stay for Black and white youth (admitted at ages 1,5,9 and 13) separated out into long-term and short-term stays for youth. Results mirror findings above, namely, that long-term stays for Black youth, and short-term stays for the youngest Black, are both longer than for the corresponding white youth.

5.4 Summary of Findings and Linking Back to Relevant Social Theory

Our empirical analysis reveals that 1) there is a relatively constant rate of entry into the NYS foster care system, 2) that youth can generally be categorized into having a *short-term* or a *long-term* stay in care, and that 3) Black youth who have long-term stays, and the youngest Black youth with short-term stays, remain on average for longer in the foster care system than their white counterparts.

In a forensic social science analysis, it is important to not only guide analyses with prior social science research, but also to tie back to this work once the analysis has been conducted. For our work, the most critical observation is the bimodal distribution in length of stay across all ages and races that differentiate youth into what we call *long-term* youth [40] and *short-term* youth [43].

Rather than social theory, however, it is practical knowledge (derived from our experiences in the system) that help us to contextualize this finding. In particular, we understand short-term stays, which are often the result of brief parental incarceration, or short-term concerns about child safety, to be a distinct class of events relative to longer-term stays, which are often the result of structural factors.

This link between long-term stays and structural factors can, in turn, help us to better understand a more novel finding from our work: differences in length of long-term stays, but not short-term stays, for Black vs. white youth. Both critical race theory in general, and critical veins of social work research, emphasize that race is a socially constructed concept intertwined with structural factors. One such factor in particular is poverty. And, although poverty is not always indicator for removal, it is often a proxy for reunification. For instance, children are not removed due to homelessness, but housing is often a requirement for a child's return home if they enter foster care [46]. Second, single motherhood, also linked to fewer financial resources, also predicts slower reunification [21]. Third, Black families have disparate access to supportive services, including employment, substance use, and mental health treatment [26]. Finally, Black youth are less likely to move from foster care into alternative plans, such as adoption and guardianship [3]. These four mechanisms are examples of how structural factors can serve to create long-term barriers to reunification for Black youth, more so than white youth.

Finally, having contextualized our findings within the literature, we turn to their policy implications. With respect to the goals of the FFPSA, and of improving life for foster youth more broadly, it is useful to distinguish between youth who are more likely to have short and long-term stays in foster care. Short term youth are typically in care for less than a month before being returned to their families. This makes them likely candidates, under the auspice of the FFPSA, to never have been removed at all [16, 43, 49]. Additionally, we note that data useful for policy in child welfare need not be highly coercive or individualized to be useful; rather, as others have noted [39], even analyses (or acknowledgement) of aggregate rates of shorter-term versus longer-term foster care stays may appraise agencies to needs related to distribution of services, and help agencies assess the success of their racial equity efforts.

In this sense, algorithms and simulations like those described below may serve as a policy function precisely via the analysis of aggregate rates and measures of allocation. To this end, we now turn to how our insights on aggregate measures from this section can be further used to construct a model that simulates the number of youth in foster care in New York state in the federal fiscal year 2018 (FY 2018), under current conditions and in response to an algorithmic intervention.

6 DATA-DRIVEN SIMULATION ANALYSIS

Informed by the discussion and findings above, we proceed to our analysis of a hypothetical algorithmic intervention in the New York state foster care system. Two aspects in particular from our forensic social science analysis carry over. First, our simulation model explicitly differentiates between youth with short-term versus long-term stays. Second, we use a variety of statistics from the work above to inform parameterization of our model. In what follows, we first provide an overview of our model which simulates the trajectory of youth within the New York state foster care system, and show it produces reliable forecasts for the number of youth in foster care. We then use this model model to investigate the potential impacts of replacing human decision-makers in the screening portion of the child welfare decision pipeline with an algorithm.

6.1 Model Overview

The goal of our model is to simulate, for each future month t in $\mathcal{T}_{future}$, the number of long- and short-term youth (denoted by l) having a select combination of race r and age a who are still in foster care system. We denote this number as $N_{r,a}^{l,t}$, and simulate it at at each stage of the decision pipeline outlined in Figure 1.

To parameterize the portion of the model where youth are in foster care, we use data from AFCARS for each month until the last month of FY 2017, i.e., each month in $\mathcal{T}_{history}$. Where data is not available to inform the model parameters (that is, when youth are not in care and thus outside the purview of the data in AFCARS), we model a range of plausible assumptions about decision pipelines of foster care system, and assess results across these possible settings. More precisely, we choose our parameters using the following steps:

- Extract the number of admitted youth $n_{r,a}^{t}$ from AFCARS, where $t \in \mathcal{T}_{history}$.
- Extracted $\mathcal{N}^{\Delta_{r,a}}$ and use it to simulate the number of admitted youth $n_{r,a}^{t}$, where $t \in \mathcal{T}_{future}$.
- Extract the long-/short-term youth proportion $\mathcal{P}_{r,a}^{l}$.
- Extract the discharge rate $\mathcal{D}_{r,a}^{l}$.

Using these parameterizations, we simulate the number of youth in care in New York using Algorithm 1. For this algorithm, we use the fact that the number of youth in foster care in the next month, $N_{r,a}^{l,t}$, equals the total number of youth who were 1) admitted into foster care during previous months and that 2) remained in the system. In turn, the number of youth who remained in the system is given by the product between the number of admitted youth in the previous month and the discharge rate (line 8-12 in Algorithm 1).

Finally, in addition to modeling how many youth are admitted, the proposed intervention we study also requires assumptions about 1) how many youth are *not* admitted (i.e. that are screened out), as well as 2) how many youth cases are screened in. Given the literature discussed above [40, 43], we assume the following:

Assumption 1. *Screen in and substantiation decisions only depend on a youth's race and whether their stay is designated as one who will have a short-term or long-term stay*

Assumption 1 says that every reported youth of a given race with a short or long-term stay will have the same probability to be screened in and substantiated. As discussed further below, this assumption notably assumes that substantiation rate does not vary with screen in rates, and vice versa. We fixed this value as introduced in Table 2. Thus, for example, given the simulated number of admitted long-term Black youth at month t, $n_{b,a}^{lo,t}$, we can obtain the number of *screened-in* long-term Black youth at month t,

Algorithm 1: Simulation

Input: $n^{t}_{r,a}$, $\mathcal{P}^{l}_{r,a}$, $\mathcal{D}^{l}_{r,a}$
Output: $N^{l,t}_{r,a}$
1 **for** $t \in \mathcal{T}_{future}$ **do**
2 **for** *l,r,a in* $\mathcal{L} \bigtimes \mathcal{R} \bigtimes A$ **do**
3 $a', t', d_t = a, t, 1$
4 $N^{l,t}_{r,a} = 0$
5 **while** $a' \geq 0$ *and* $t' \geq$ *Nov2000* **do**
6 $\widehat{\Delta}^{l,t}_{r,a} = \mathcal{P}^{l}_{r,a'} \cdot n^{t'}_{r,a'}$
7 // # of admitted long/short-term youth
8 $\widehat{\Delta}^{l,t}_{r,a} = \widehat{\Delta}^{l,t}_{r,a} \cdot \mathcal{D}^{l}_{r,a}(d_t)$
 // # of remaining long/short-term youth after d_t months
9 $N^{l,t}_{r,a}$ += $\widehat{\Delta}^{l,t}_{r,a}$
10 $d_t = d_t + 1$
11 $t' = t' - 1$
12 **if** $(d_t - 6) | 12$ **then**
13 $a' = a' - 1$

Notations:
r : Race of youth
a : Age of youth
l : Long and short-term membership of youth
$n^{t}_{r,a}$: The number of admitted youth at month t
$N^{l,t}_{r,a}$: The number of remaining youth of at month t
$\mathcal{P}^{l}_{r,a}$: Percentage of long-/short-term youth of total admitted youth
$\mathcal{D}^{l}_{r,a}$: The discharge rate

$\dot{n}^{lo,t}_{b,a} = \frac{n^{lo,t}_{b,a}}{R^{sub,lo}_{b}}$, and the number of *reported* long-term Black youth at month t, $\ddot{n}^{lo,t}_{b,a} = \frac{n^{lo,t}_{b,a}}{R^{sub,lo}_{b} \cdot R^{scr,lo}_{b}}$.

6.2 Validation of the Base Model

Figure 4 shows that, given data from 2017 and before, our model generates forecasts for the number of youth in foster care in New York state in 2018 that are in line with real data. The figure shows model predictions (triangle points) for Black (red) and white (cyan) youth separately, comparing simulated estimates aggregated over long-term and short-term youth of all ages. It compares these predictions to the ground truth values (circle points) in the AFCARS dataset. The Pearson correlation between the number of simulated Black youth and the number of actual Black youth in foster care is 0.98 ($p < 0.001$). The Pearson correlation between the number of simulated white youth and the number of real white youth in foster care is also 0.98 ($p < 0.001$). And the Pearson correlation between the proportion of Black youth in simulated foster care system and the proportion of Black youth in the actual data is 0.78 ($p = 0.004$). The results show that our simulation model is a reliable starting point from which to model potential interventions resulting from the implementation of the FFPSA in 2018.

6.3 Extending the Model to Investigate the Effects of an Algorithmic Intervention

Our hypothetical intervention is based on the work of Chouldechova et al. [8], who develop a model to assist screening decisions in Allegheny County, PA. We stress that Chouldechova et al. [8] do not intend for their model to replace humans, and that there are important effects of maintaining a human in the loop during this decision process [7]. However, it is nonetheless informative to study the simplified case where the model does, in fact, fit this role. Below, we detail our (simulated) implementation of their algorithm, and the additional model parameters we vary to explore how the model responds to different potential context in which it is deployed.

6.3.1 Simulating Training and Deployment of the Algorithm. To build the machine learning algorithm to identify the youth who need to be screened in, Chouldechova et al. [8] use what we will call the *profiles* of youth. These profiles include demographics, behavioral health records and past history in care (among other variables) as features, and well as whether or not the youth end up in the foster care as the outcome to be predicted.

To simulate the training procedure for their model, we define (and simplify) the *features* given to the model to be continuous values drawn from a Gaussian distribution. Given our knowledge above that youth with short- and long-term stays are likely to have distinct reasons for being brought into care, we further assume that their profiles are generated from two different distributions. Profiles of long-term youth are generated from $\mathcal{N}^{lo}(\mu_{lo}, \sigma_{lo})$ while profiles of short-term youth are generated from $\mathcal{N}^{sh}(\mu_{sh}, \sigma_{sh})$. Parameters of these two distributions are shown in Table 2.

We define the *label* the model is trained on using the (simulated) decisions at the substantiation phase. Notably, as shown in Figure 1, Chouldechova et al.'s 2018 model makes predictions at the screen-in stage of the child welfare pipeline. Thus, the label used is a downstream proxy assumed to be less racially biased. In this setup, however, [8] therefore assume that if youth end up in foster care at the last decision made in the decision pipeline, then the youth should be screened in at the first decision, and vice versa.

We simulate the impact of introducing the model from Chouldechova et al. [8] as if it was trained using a dataset generated before November 2017, and implemented during the following year. To do so, we construct the training set by using profiles of admitted youth as positive samples and using profiles of reported but not admitted youth as negative samples. After the training dataset is constructed, we use it to train a logistic regression model as a screen-in recommendation tool. [3] In our simulation, we model the deployment of this machine learning model as having complete autonomy over decision-making. To be more specific, after deploying the algorithm, step (8) in Algorithm 1 is changed to send the profiles of $\ddot{n}^{l,t}_{r,a}$ reported youth at each month to the deployed algorithms to extract youth screened in. And then, admitted youth are extracted randomly using the substantiate rate introduced in the Table 2 following Assumption 1, because substantiation decisions are still made entirely by humans.

6.3.2 Varying Assumptions to Evaluate the Intervention. We evaluate the performance of this algorithmic intervention under different assumptions about the underlying foster care system on which it is trained. These variations are summarized in Table 2 and detailed below.

Separability of Short- vs. Long-term Youth Profiles. We model the difficulty of distinguishing between long-term foster care youth and short-term foster care youth from their profiles as sampling

[3] Training data was balanced by down-sampling negative samples to approximate how machine learning models are trained in real world. Note that we don't modify hyperparameters of the model (e.g. the decision threshold) because this induces yet another implicit value judgement which is out of the scope of our main focus in the current work.

Table 2: Tabular description of parameters involved in the construction of our simulation model. The red value of the parameter is the default value

Parameter	Values Taken	Parameter	Values Taken
$R_b^{scr,lo}$: Screen in rate of long-term Black youth	70%	$R_b^{scr,sh}$: Screen in rate of short-term Black youth	6%,8%,10%
$R_w^{scr,lo}$: Screen in rate of long-term white youth	70%	$R_w^{scr,sh}$: Screen in rate of short-term white youth	10%
$R_b^{sub,lo}$: Substantiate rate of long-term Black youth	70%	$R_b^{sub,sh}$: Substantiate rate of short-term Black youth	10%
$R_w^{sub,lo}$: Substantiate rate of long-term white youth	70%	$R_w^{sub,sh}$: Substantiate rate of short-term white youth	10%
$\mathcal{N}^{lo}(\mu_{lo}, \sigma_{lo})$: Distribution of feature of long-term youth	([1, 1], [2, 2], [10, 10], [[1, 0][0, 1]])	$\mathcal{N}^{sh}(\mu_{sh}, \sigma_{sh})$: Distribution of feature of short-term youth	([0, 0], [[1, 0][0, 1]])

long-term foster care youth's profile from different Gaussian distributions. For example, when long-term foster care youth's profiles are sampled from $\mathcal{N} \sim ([10, 10], [[1, 0], [0, 1]])$ and short-term foster care youth's profiles are sampled from $\mathcal{N} \sim ([0, 0], [[1, 0], [0, 1]])$, it is much easier for an algorithm to differentiate them, compared to the situation where long-term foster care youth's profiles are sampled from $\mathcal{N} \sim ([1, 1], [[1, 0], [0, 1]])$ and short-term foster care youth's profiles are sampled from $\mathcal{N} \sim ([0, 0], [[1, 0], [0, 1]])$. We alter the mean of the feature distribution of long-term foster care youth in $([10, 10], [2, 2], [1, 1])$ to represent *high* ,*moderate* and *low* separability between youth who will have long versus short term stays, leaving other parameters at their defaults.

Modeling Racial Biases in Reporting. We model reporting bias across racial lines in the current foster care system by varying the screen-in rates of short-term Black youth. Controlling for the number of admitted Black youth with short-term stays, the lower the screen-in rate, $R_b^{scr,sh}$, for Black youth with short-term stays, the larger the number of Black youth reported. We vary the screen in rate for short-term Black youth in (6%,8%,10%) to represent *high reporting bias*, *low reporting bias*, and *no reporting bias*, leaving other parameters at their defaults.

6.4 Results of Assessing an Algorithmic Intervention

Overall, we find that relying on an algorithm trained as suggested by Chouldechova et al. [8] to make screen-in decision would increase the number of youth in foster care, contradicting the goals of the FFPSA. However, the impact of the algorithm varies depending on the assumed separability between profiles of youth with short-term vs. long-term stays. Figure 5 compares monthly forecasts of the number of Black and white youth in foster care with and without using the algorithm for making screen-in decisions. Moreover, we distinguish between different algorithm-based screen-in decisions by the ease of separability between youth with short- and long-term stays. We see that in all conditions of separability we examined, the algorithm would increase the number of youth in foster care. However, perhaps surprisingly, the increase is significantly more pronounced when it is easy to separate long- from short- term youth. That is, the greater the assumed differences between youth with long and short-term stays, the more youth the algorithm puts into care.

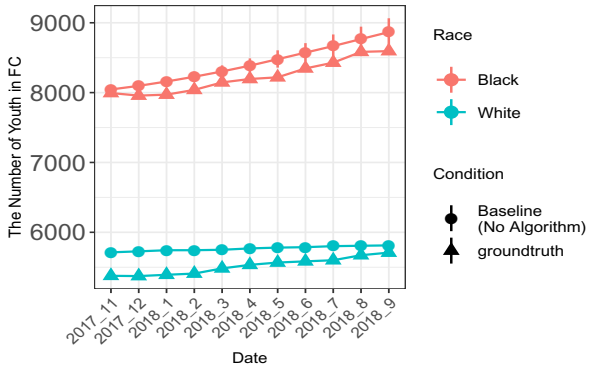


Figure 4: The number (y-axis) of Black (red line) and white (turqouise line) youth in foster care, by month (x-axis), for both the simulated (circles) and real (triangle) foster care data in FY 2018. Real data is derived from the AFCARS dataset for New York state, simulated data from our general simulation model.

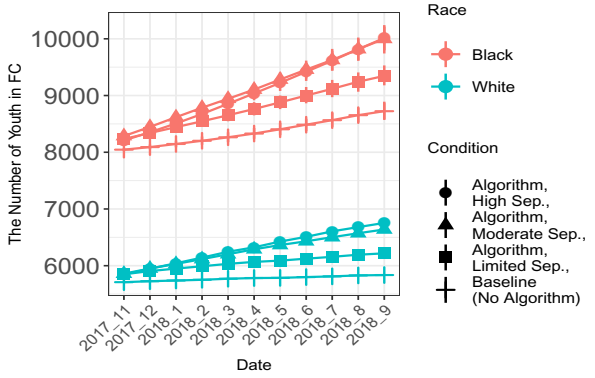
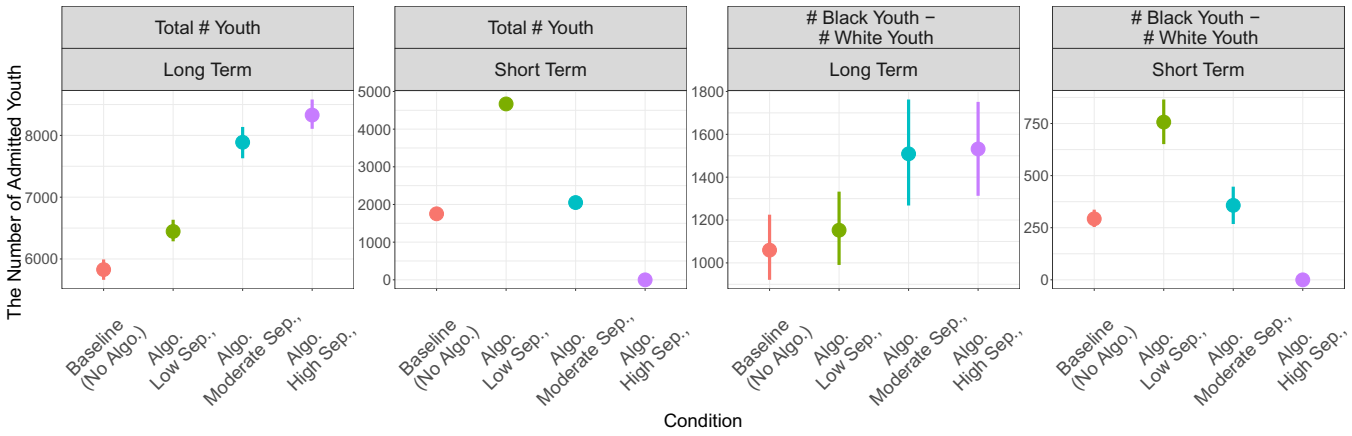
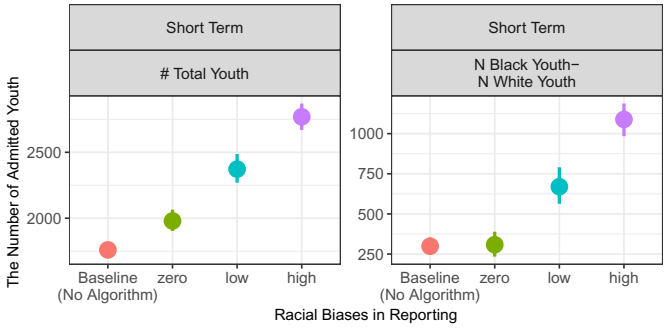


Figure 5: The simulated number of Black and white youth in foster care by month for different screen-in procedures: without an algorithm and with an algorithm but different levels of separability between short-term and long-term youth profiles.

Figure 6a shows a) that aggregating over youth with long- and short-term stays provides a distorted picture of the algorithm's benefits, and b) that, depending on the separability assumptions, the



(a) The number of admitted long-term and short-term youth during simulated months depending on the difficulty of separability. The first two plots shows the total number of long- and short-term stays while the last two plots shows the disparities between the number of Black and white youth for long- and short-term stays.



(b) The number of admitted short-term youth during simulated months for different levels of racial biases in reporting in the original system. The first plot shows the total number of short term youth while the second plot shows the disparity of short term Black and white youth.

Figure 6: Impact of the algorithm under varying assumptions

algorithm has non-obvious side-effects on racial disparities. More specifically, as the separability of youth with long- and short-term stays increases, the number of youth with long-term stays in foster care increases (left most plot), and so does the racial inequity for these youth (right middle plot). However, both the number of youth with short-term stays (left middle plot) *and* the racial inequity (right most plot) in the number of youth with short-term stays increase as separability decreases. Thus, under the assumptions of our model, when training an algorithm under a situation where youth who end up in care long-term are similar (as far as the algorithm is concerned) to youth who stay for a short period of time, the *more* short term youth enter foster care overall. This is because the screen-in algorithm, if trained using substantiation as a proxy variable, will get better at identifying youth who the current system sees as needing long-term care, and these youth will, in turn, remain in care for longer. Even under a seemingly better situation where there is high separability between profiles of long-term and short-term youth, algorithms will place more long-term youth into the system, contradicting the goal of FFPSA.

According to our simulation, higher levels of bias when reporting youth to CPS result in more youth in care and larger racial disparities (see Figure 6b). The explanation for this phenomenon comes from the interplay of the different parts of the system: if short-term youth are more often Black, then while screen-in rates of Black youth will decrease overall, the algorithm will produce more false positives for Black youth. This effect is problematic since the reporting of youth is (a) biased [22] and (b) exogenous to the child welfare system, so a solution to this issue is needed and difficult to implement.

7 DISCUSSION AND CONCLUSION

Two major stated goals of the FFPSA are to 1) reduce unnecessary admission of youth into the foster care system and 2) provide more family-based service for youth who are in the foster care system. To achieve such goals, practitioners have been incentivized to leverage additional data collection and predictive modeling to assist decision making [1].

Here, we construct a simulation model that considers how a particular algorithm can help to address the goals of the FFPSA, while

maintaining the additional goal of racial equality. Before reviewing the main resultant claims, we find it pertinent to emphasize their limits. First, as noted above, the intervention we consider does not account for the fact that while this may not be true of tomorrow's algorithms [36], today's algorithms are often used, at best, as suggestions for screening in, rather than all-encompassing decision-makers [8]. Second, and related, our model greatly simplifies the ways in which algorithmic decisions are made about specific youth in foster care. Additional complexity, e.g. by creating more realistic youth profiles using advanced machine learning models, may lead to distinct conclusions. Third, certain parameterizations may be reasonably disagreed with. In particular, it is possible that higher levels of screen-in rates may lead directly to lower rates of substantiation, which violates Assumption 1.[4] As such, we expect, and hope, that our model is viewed primarily as a starting point for informed debate.

To this end, and with these limitations in mind, we use our model here to bring forth the point that implementing a machine learning algorithm to assist in screen-in decisions will tend instead to increase the number of youth in foster care, and to increase racial disparities in the number of Black versus white youth as well. Importantly, this increase of youth in care often comes in the form of better identification of youth that may be in genuine need of long-term foster care *as defined by the current system*. In turn, while increasing the total number of youth in care, the algorithm actually decreases the number of and racial disparities in short-term stays in foster care (i.e. youth who may be candidates for deferral from foster care through prevention services). As we interpret the work of Chouldechova et al. [8], the algorithm in this sense is accomplishing its goal—reducing racial inequality and mis-identified screen in cases *as determined by the current system*.

Our findings thus expose a contradiction between the two justifiable goals of FFPSA: 1) placing fewer youth in foster care and 2) more accurate identification of families who need more intensive services according to current measures. Namely, improving our ability to differentiate youth with long-term versus short-term stays in the current system increases the number of overall youth in foster care, but does so through the inclusion of youth who the system would likely deem to need care. Therefore, we argue that implementing algorithms that use past historical data under the new directive of the FFPSA will therefore require either 1) that we embed different notions of who needs foster care into our models or 2) that we use variables other than substantiation rates to train them.

Equally as important, our simulation (and other analyses of the child welfare system [1]) make clear that efforts to develop algorithms for screening in youth must provide space for discussion about who, if anyone, *should* be placed into foster care. Proposed algorithms shouldn't be necessarily fitted into the existing problematic loop of the social system but rather provide a new perspective to policy makers that might facilitate more effective decision-making about policy and practice at the macro level.

Our work also shows that these decisions must occur with a deeper consideration of assumptions we make about the myriad and sometimes unknowable parameters of the existing system. At a fundamental level, our work shows that the assumption of substantiation decisions as a proxy for screen-in decisions made by Chouldechova et al. [8] can be problematic. Less obvious, however, is that variations in assumptions about parts of the system external to this proxy can still have important effects on model outcomes. This variation in our simulation model not simply hypothetical. Wildly different policy environments across, and even within, states mean that assumptions in one setting are quite possibly incorrect for others. These complex and hierarchical sets of assumptions benefit from the ability to systematically investigate how one change may impact others in non-obvious ways. Our simulation tool, we therefore hope, can serve as a kind of test-bed as more data becomes available to validate or inform certain assumptions in different jurisdictions, including New York state. We also hope it will assist practitioners and policy-makers in understanding the trade-offs implicit in different assumptions they are making about decisions made in their own local arena.

[4]We thank our reviewers for pointing out this particular issue

ACKNOWLEDGMENTS

This research was supported by an NSF award IIS1939579, with partial support from Amazon. The NSF and Amazon had no role in the design and conduct of the study; access and collection of data; analysis and interpretation of data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication. The authors declare no other financial interests. The authors would, however, like to thank members of the UB FAI team for thoughts on this article, as well as FAccT and EAAMO reviewers who evaluated previous versions of the paper.

REFERENCES

[1] J Khadijah Abdurahman. 2021. Calculating the Souls of Black Folk: Predictive Analytics in the New York City Administration for Children's Services. *Columbia Journal of Race and Law* 11, 4 (2021), 75–110.

[2] Rediet Abebe, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G. Robinson. 2020. Roles for Computing in Social Change. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 252–260. https://doi.org/10.1145/3351095.3372871

[3] Becci A Akin. 2011. Predictors of foster care exits to permanency: A competing risks analysis of reunification, guardianship, and adoption. *Children and Youth Services Review* 33, 6 (2011), 999–1011.

[4] Marion A. Becker, Neil Jordan, and Rebecca Larsen. 2007. Predictors of Successful Permanency Planning and Length of Stay in Foster Care: The Role of Race, Diagnosis and Place of Residence. *Children and Youth Services Review* 29, 8 (Aug. 2007), 1102–1113. https://doi.org/10.1016/j.childyouth.2007.04.009

[5] Reiko Boyd. 2014. African American disproportionality and disparity in child welfare: Toward a comprehensive conceptual framework. *Children and Youth Services Review* 37 (2014), 15–27.

[6] Children's Bureau. 2014. Issue Brief: In-Home Services in Child Welfare. https://www.childwelfare.gov/pubPDFs/inhome_services.pdf

[7] Hao-Fei Cheng, Logan Stapleton, Ruiqi Wang, Paige Bullock, Alexandra Chouldechova, Zhiwei Steven Steven Wu, and Haiyi Zhu. 2021. Soliciting Stakeholders' Fairness Notions in Child Maltreatment Predictive Systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 390, 17 pages. https://doi.org/10.1145/3411764.3445308

[8] Alexandra Chouldechova, Diana Benavides-Prado, Oleksandr Fialko, and Rhema Vaithianathan. 2018. A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Proceedings of Machine Learning Research, Vol. 81)*, Sorelle A. Friedler and Christo Wilson (Eds.). PMLR, New York, NY, USA, 134–148.

[9] Alexander D'Amour, Hansa Srinivasan, James Atwood, Pallavi Baljekar, D. Sculley, and Yoni Halpern. 2020. Fairness is Not Static: Deeper Understanding of Long Term Fairness via Simulation Studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain)

(FAT '20).* Association for Computing Machinery, New York, NY, USA, 525–534. https://doi.org/10.1145/3351095.3372878

[10] Maria De-Arteaga, Riccardo Fogliato, and Alexandra Chouldechova. 2020. *A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores.* Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376638

[11] Alan J Dettlaff, Kristen Weber, Maya Pendleton, Reiko Boyd, Bill Bettencourt, and Leonard Burton. 2020. It is not a broken system, it is a system that needs to be broken: The upEND movement to abolish the child welfare system. *Journal of Public Child Welfare* 14, 5 (2020), 500–517.

[12] David A Dickey and Wayne A Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association* 74, 366a (1979), 427–431.

[13] Joshua M. Epstein. 2008. Why model? *Journal of Artificial Societies and Social Simulation* 11, 4 (2008), 12.

[14] NCCPR: National Coalition for Child Protection Reform. 2019. No, you can't use predictive analytics to reduce racial bias in child welfare. https://www.nccprblog.org/2019/06/no-you-cant-use-predictive-analytics-to.html

[15] Patrick J Fowler, Katherine E Marcal, Saras Chung, Derek S Brown, Melissa Jonson-Reid, and Peter S Hovmand. 2020. Scaling up housing services within the child welfare system: policy insights from simulation modeling. *Child maltreatment* 25, 1 (2020), 51–60.

[16] Antonio Garcia, Alan Puckett, Mark Ezell, Peter J Pecora, Theresa Tanoury, and Wenceslao Rodriguez. 2014. Three models of collaborative child protection: what is their influence on short stays in foster care? *Child & Family Social Work* 19, 2 (2014), 125–135.

[17] Nigel Gilbert and Klaus G Troitzsch. 2005. *Simulation for the Social Scientist.* Open University Press, USA.

[18] Jeremy Goldhaber-Fiebert and Lea Prince. 2019. Impact Evaluation Summary of the Allegheny Family Screening Tool.

[19] Jeremy D Goldhaber-Fiebert, Stephanie L Bailey, Michael S Hurlburt, Jinjin Zhang, Lonnie R Snowden, Fred Wulczyn, John Landsverk, and Sarah M Horwitz. 2012. Evaluating child welfare policies with decision-analytic simulation models. *Administration and Policy in Mental Health and Mental Health Services Research* 39, 6 (2012), 466–477.

[20] Ben Green. 2021. The Flaws of Policies Requiring Human Oversight of Government Algorithms.

[21] Marian S Harris and Mark E Courtney. 2003. The interaction of race, ethnicity, and family structure with respect to the timing of family reunification. *Children and Youth Services Review* 25, 5-6 (2003), 409–429.

[22] Marian S Harris and Wanda Hackett. 2008. Decision points in child welfare: An action research model to address disproportionality. *Children and Youth Services Review* 30, 2 (2008), 199–215.

[23] Lily Hu, Nicole Immorlica, and Jennifer Wortman Vaughan. 2019. The Disparate Effects of Strategic Manipulation. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) *(FAT* '19).* Association for Computing Machinery, New York, NY, USA, 259–268. https://doi.org/10.1145/3287560.3287597

[24] Stefania Ionescu, Anikó Hannák, and Kenneth Joseph. 2021. An Agent-Based Model to Evaluate Interventions on Online Dating Platforms to Decrease Racial Homogamy. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21).* Association for Computing Machinery, New York, NY, USA, 412–423. https://doi.org/10.1145/3442188.3445904

[25] Emily Keddell. 2019. Algorithmic justice in child protection: Statistical fairness, social justice and the implications for practice. *Social Sciences* 8, 10 (2019), 281.

[26] Theresa Knott and Sulaimon Giwa. 2012. African American disproportionality within CPS and disparate access to support services: Review and critical analysis of the literature. *Residential Treatment for Children & Youth* 29, 3 (2012), 219–230.

[27] David J Kolko, Michael S Hurlburt, Jinjin Zhang, Richard P Barth, Laurel K Leslie, and Barbara J Burns. 2010. Posttraumatic stress symptoms in children and adolescents referred for child welfare investigation: A national sample of in-home and out-of-home care. *Child maltreatment* 15, 1 (2010), 48–63.

[28] David Lazer, Alex Sandy Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, et al. 2009. Computational Social Science. *Science (New York, NY)* 323, 5915 (2009), 721.

[29] Bethany R Lee and Ron Thompson. 2008. Comparing outcomes for youth in treatment foster care and family-style group care. *Children and Youth Services Review* 30, 7 (2008), 746–757.

[30] Donald Martin, Vinodkumar Prabhakaran, Jill A. Kuhlberg, A. Smart, and William S. Isaac. 2020. Extending the Machine Learning Abstraction Boundary: A Complex Systems Approach to Incorporate Societal Context.

[31] Daniel A. McFarland, Kevin Lewis, and Amir Goldberg. 2015. Sociology in the Era of Big Data: The Ascent of Forensic Social Science. *The American Sociologist* 47, 1 (2015), 1–24.

[32] Amy McKlindon. 2019. Applying the research and evaluation provisions of the Family First Prevention Services Act.

[33] Smitha Milli, John Miller, Anca D. Dragan, and Moritz Hardt. 2019. The Social Cost of Strategic Classification. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) *(FAT* '19).* Association for Computing Machinery, New York, NY, USA, 230–239. https://doi.org/10.1145/3287560.3287576

[34] US Department of Health, Human Services, et al. 2019. Adoption and Foster Care Analysis and Reporting System (AFCARS) FY 2000-2018 data.

[35] Nathanael J. Okpych. 2015. Receipt of Independent Living Services among Older Youth in Foster Care: An Analysis of National Data from the U.S. *Children and Youth Services Review* 51 (April 2015), 74–86. https://doi.org/10.1016/j.childyouth.2015.01.021

[36] Joanna Redden, Lina Dencik, and Harry Warne. 2020. Datafied child welfare services: unpacking politics, economics and power. *Policy Studies* 0, 0 (Feb. 2020), 1–20.

[37] Horst WJ Rittel and Melvin M Webber. 1973. Dilemmas in a general theory of planning. *Policy sciences* 4, 2 (1973), 155–169.

[38] Dorothy E Roberts. 2019. Digitizing the carceral state.

[39] Maria Y Rodriguez, Diane DePanfilis, and Paul Lanier. 2019. Bridging the gap: Social work insights for ethical algorithmic decision-making in human services. *IBM Journal of Research and Development* 63, 4/5 (2019), 8–1.

[40] Rachel Rosenberg and Samuel Abbott. 2019. Supporting Older Youth Beyond Age 18: Examining Data and Trends in Extended Foster Care. https://www.childtrends.org/publications/supporting-older-youth-beyond-age-18-examining-data-and-trends-in-extended-foster-care

[41] Jesse Russell and Stephanie Macgill. 2015. Demographics, Policy, and Foster Care Rates; A Predictive Analytics Approach. *Children and Youth Services Review* 58 (Nov. 2015), 118–126. https://doi.org/10.1016/j.childyouth.2015.09.009

[42] Anjana Samant, Aaron Horowitz, Sophie Beiers, and Kath Xu. 2021. *Family Surveillance by Algorithm: The Rapidly Spreading Tools Few Have Heard Of.* ACLU. https://www.aclu.org/news/womens-rights/family-surveillance-by-algorithm-the-rapidly-spreading-tools-few-have-heard-of

[43] Vivek S Sankaran and Christopher Church. 2016. Easy come, easy go: The plight of children who spend less than thirty days in foster care. *U. Pa. JL & Soc. Change* 19 (2016), 207.

[44] Devansh Saxena, Karla Badillo-Urquiola, Pamela J. Wisniewski, and Shion Guha. 2020. *A Human-Centered Review of Algorithms Used within the U.S. Child Welfare System.* Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376229

[45] Thomas C. Schelling. 1971. Dynamic models of segregation. *The Journal of Mathematical Sociology* 1, 2 (1971), 143–186.

[46] Marybeth Shinn, Jessica Gibbons-Benton, and Scott R Brown. 2015. Poverty, homelessness, and family break-up. *Child welfare* 94, 1 (2015), 105.

[47] J. Sterman. 2002. System Dynamics Modeling: Tools for Learning in a Complex World. *California Management Review* 43 (2002), 25 – 8.

[48] Yvonne A Unrau, John R Seita, and Kristin S Putney. 2008. Former foster youth remember multiple placement moves: A journey of loss and hope. *Children and Youth Services Review* 30, 11 (2008), 1256–1266.

[49] Eric van Santen. 2010. Predictors of exit type and length of stay in non-kinship family foster care—The German experience. *Children and Youth Services Review* 32, 10 (2010), 1211–1222.

[50] Elizabeth Weigensberg, Nuzhat Islam, et al. 2020. *SCAN Policies Database Fact Sheet.* Technical Report. Mathematica Policy Research.

[51] Fred Wulczyn and Bridgette Lery. 2007. Racial disparity in foster care admissions.

[52] Jason Yan, Melanie Sage, Seventy F. Hall, Yuhao Du, and Kenneth Joseph. 2021. A Computational Social Science Approach to Understanding Predictors of Chafee Service Receipt. arXiv:2111.14901 [cs.SI]

[53] Anna Yelick and Bruce Thyer. 2020. The effects of family structure and race on decision-making in child welfare. *Journal of Public Child Welfare* 14, 3 (May 2020), 336–356.