WA IV-E
Family Assessment Response

Description of the Outcome Analysis (Companion to Family Assessment Response Interim Evaluation Report)

July 1, 2018

TriWest
Health + Human Service Evaluation + Consulting
4450 Arapahoe Avenue, Suite 100, Boulder, CO 80303

Table of Contents

Overview

This document is a companion to the 2018 Interim Evaluation Report and describes the structure of the data, steps used in their analysis, and analysis results for the Washington State FAR implementation. It includes technical descriptions of data manipulation and statistical analysis, and it is written for audiences with expertise in propensity score matching, multiple regression, and other econometric techniques used in program analysis. It is also part of the TriWest Group (TriWest) peer-review process, by which individuals familiar with Children's Administration (CA) data may verify the appropriate use of those data. All analysis is done in the statistical computing software, R, and R scripts are available to reviewers desiring additional detail on the steps in data manipulation and analysis.

FAR Program Overview and Implementation

Washington State's Title IV-E Waiver demonstration project is an implementation of Family Assessment Response (FAR), a differential response pathway for screened-in allegations of abuse and neglect as an alternative to traditional investigations. FAR's goal was to produce superior child welfare outcomes by providing support to families subject to claims of child neglect or minor abuse, rather than directing them to child welfare investigations. Support has taken several forms, including the direct assistance of Children's Administration social workers, evidence-based treatments provided by agencies contracted by the Children's Administration, connection to local community resources, and provision of concrete goods and services.

FAR, which focuses on determining whether child abuse or neglect occurred and whether children are subject to ongoing safety threats, is an alternative to a child welfare investigation. The hypothesis underlying FAR is that providing assistance in times of crisis may avert more serious consequences such as placement of children in foster care and future child abuse incidents. The FAR implementation created two pathways for responding to abuse and neglect allegations: (1) FAR for lower risk cases and (2) child welfare investigations for cases with greater risk. As offices implemented FAR, separate administrative units were created for FAR and investigative social workers.

Washington State child welfare offices implemented FAR at different times. Three of the 47 offices implemented FAR in January 2014. Six more implemented in July 2014, with additional offices rolling out each quarter, culminating in the complete statewide rollout in June 2017, with the FAR model active in all 47 offices.

Description of the Data Generating Process

Selecting Data Groups for Comparisons

Starting in January 2014, all families subject to claims of abuse and neglect (intakes) were evaluated for eligibility for the FAR pathway, both in offices that implemented FAR and in offices in which FAR had not yet been implemented. During the staggered implementation, some families subject to accusations of child abuse or neglect (intakes) may or may not have received FAR, depending on their location. The presence of both FAR and “FAR-eligible” families (i.e., families who likely would have been assigned FAR if their local office had been implemented) drives the core of TriWest’s data analysis plan: it allows comparison of outcomes between families receiving FAR in FAR-implemented offices (i.e., treatment group) to FAR-eligible families subject to investigation in offices that had not yet implemented FAR (i.e., comparison group). Families excluded from FAR involved accusations of a more serious nature with significantly greater risk of child harm. These families were automatically assigned to the investigative pathway and were not used in our analysis.

The staggered rollout of FAR resulted in many more comparison families than FAR families in the January 2014 cohort (explained below): 8,515 to 663, after various filters were applied. By the second cohort, starting in July 2014, the balance was 4,953 to 2,602. In the third and fourth cohorts (January and July 2015, respectively) there were more FAR families than comparison families, and we therefore randomly selected 2,000 FAR families for Cohort 3 and 1,000 FAR families for Cohort 4. By reducing the size of the FAR cohort to a number below the FAR-eligible investigative cohort, we created the opportunity to match the remaining families to FAR-eligible investigative families with similar characteristics. By the time we considered the seventh and final cohort, starting in January 2017, the low number of available comparison group families resulted in a matching of 250 FAR and 250 FAR-eligible families.

Not all families assigned to FAR remained on that pathway. FAR is a voluntary program, and families that declined FAR may have been transferred (involuntarily) to investigations. This analysis is an “intent-to-treat” design: once a family was screened into FAR, including any immediate supervisor overrides, the family was considered a FAR family for the duration of the analysis. We performed a separate analysis (not reported in this document) comparing families that completed FAR versus those that were transferred to investigation. Not surprisingly, outcomes were much better for those families remaining in the FAR program.

The following figure represents the number of intakes received in all seven cohorts, broken down by screening decision (Screened Out, FAR, Investigation, Risk Only, Missing Values), and further subdivided by FAR eligibility and eventual disposition.

Cohort Sample Periods			
Cohort 1: Jan–Jun, 2014	Cohort 3: Jan–Jun, 2015	Cohort 5: Jan–Jun, 2016	Cohort 7: Jan–Jun, 2017
Cohort 2: Jul–Dec, 2014	Cohort 4: Jul–Dec, 2015	Cohort 6: Jul–Dec, 2016	

Cases Screened Out (Intake type=0)	
Cohort 1:	12,035
Cohort 2:	10,197
Cohort 3:	9,984
Cohort 4:	8,251
Cohort 5:	9,129
Cohort 6:	7,945
Cohort 7:	8,832
Totals	**66,373**

Total Intakes	
Cohort 1:	25,566
Cohort 2:	21,277
Cohort 3:	22,206
Cohort 4:	19,245
Cohort 5:	20,496
Cohort 6:	17,725
Cohort 7:	20,119
Totals	**146,634**

Missing Values (Intake type=NA)	
Cohort 1:	3
Cohort 2:	75
Cohort 3:	299
Cohort 4:	328
Cohort 5:	313
Cohort 6:	327
Cohort 7:	256
Totals	**1,601**

Risk-Only Cases (Intake type=3)	
Cohort 1:	1,077
Cohort 2:	996
Cohort 3:	901
Cohort 4:	1,045
Cohort 5:	986
Cohort 6:	1,061
Cohort 7:	1,178
Totals	**7,244**

FAR Cases (Intake type=1)	
Cohort 1:	664
Cohort 2:	2,629
Cohort 3:	5,589
Cohort 4:	5,429
Cohort 5:	5,934
Cohort 6:	5,473
Cohort 7:	7,172
Totals	**32,890**

Investigative Cases (Intake type=2)	
Cohort 1:	11,787
Cohort 2:	7,380
Cohort 3:	5,433
Cohort 4:	4,192
Cohort 5:	4,134
Cohort 6:	2,919
Cohort 7:	2,681
Totals	**38,526**

FAR Case Disposition (of 8,897)	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 6	Cohort 7	Total
0=Missing	0	0	0	3	0	3	146	152
1=Remained FAR	597	2,328	4,905	4,823	5,262	4,889	6,222	29,026
2=Declined FAR	39	170	315	292	298	286	322	1,722
3=Transferred (including investigation)	27	80	124	125	140	130	230	856

Case That Would Have Been Eligible for FAR If Available (Potential Comparison Observations)	
Cohort 1:	9,152
Cohort 2:	5,378
Cohort 3:	3,277
Cohort 4:	2,014
Cohort 5:	1,936
Cohort 6:	1,104
Cohort 7:	556
Totals	**23,427**

Cases Not Eligible for FAR Even If Available	
Cohort 1:	2,551
Cohort 2:	1,920
Cohort 3:	2,075
Cohort 4:	2,127
Cohort 5:	2,142
Cohort 6:	1,775
Cohort 7:	2,092
Totals	**14,682**

Investigative Cases Marked Eligible and Emergent	
Cohort 1:	84
Cohort 2:	82
Cohort 3:	81
Cohort 4:	51
Cohort 5:	56
Cohort 6:	40
Cohort 7:	23
Totals	**417**

TriWest
Health + Human Service Evaluation + Consulting

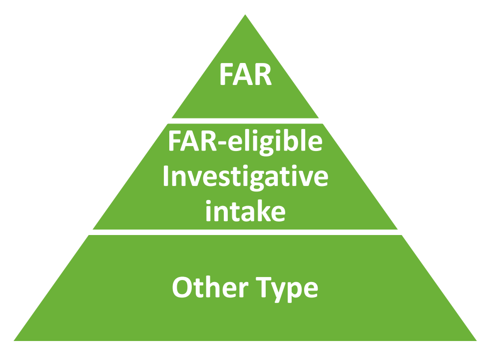
Data Sets and Significant Variables

The processing of administrative data sets occurred as each new cohort became available. We received separate six-month data files from Washington State. Each data transfer included files of two types: (1) a single file of pre-existing characteristics for each family in the new cohort (the cohort file) and (2) files of outcome variables for families in the most recent and all previous cohorts. The outcome variables represent events subsequent to each family's intake (e.g., new intakes, child removals, or services received). The cohort files were static in the sense that all information included was drawn from the events before the family's intake.

In these data sets, each row represents a single family, each of which was coded using the variable *Intktype* as Screened Out (not accepted), FAR, Investigative, or Risk Only (a category representing abuse or neglect claims that do not display sufficient risk to warrant CA intervention). Both the cohort and outcome files drew from both FamLink (CA's data management system) and other Washington State data systems related to criminal justice, economic assistance, mental health, physical health, and other social service systems.

Within the data, families were identified with the numeric variable *ID_CASE*. Because an ID_CASE may have multiple intakes during a cohort period, and the intake type (FAR, Investigative, Risk Only) may vary with each new intake, we categorized a given family during a cohort period with the following prioritization: actual FAR, FAR-eligible investigative intakes, and all other intake types. As an example, if a family's first intake was Risk Only, and one month later the family had a FAR intake, the family was categorized as FAR within that cohort period since actual FAR is prioritized over Risk Only intakes.

Intake Categorization Priority



This prioritization also applied to families that had intakes in multiple periods. If a family had a FAR intake in Cohort 6, the process that the Washington State Research and Data Analysis (RDA) team used to generate the cohorts removed that family's Investigative and Risk Only intakes from earlier cohorts. For this reason, the table of intakes reported on the previous page does not represent the total number of intakes for all families; it is instead the unduplicated count by ID_CASE of intakes during the seven cohort periods.

Because of this prioritization of FAR intakes, TriWest's data did not include all intakes. Since the purpose of the analysis is to measure the effect of FAR, this limitation did not impact our analysis of the comparison of FAR (treatment) to matched FAR-eligible investigative (comparison) intakes.

Disparity

For issues related to disparity in screening, all intakes are required, and we have requested a supplemental data set that includes all intakes. We will report on disparity in the final evaluation report.

Selecting Outcome Variables

TriWest focuses on outcomes addressing the following research questions:

- **Removals.** Does FAR reduce the number of children removed from their families?
- **Re-Referrals.** Does FAR reduce future accusations of abuse and neglect?
- **Costs.** Does FAR reduce the costs to CA of serving families?

Removals

To address the first research question, we used the outcome files to generate a series of binary outcome variables indicating whether a family had one or more children removed during the specified time period. Time periods included spans within 3 months (90 days), 6 months, 12 months, and 24 months of intake. Variable names were *removal3*, *removal6*, *removal12*, and *removal24*. These are binary indicator variables; they did not capture how many children were removed from a family, but only whether a family experienced one or more removals. Because of complexities in identifying the unduplicated count of unique children that were removed, we are more confident in a binary measure for removals.

Re-Referrals

To address the second research question, we created binary variables indicating whether a family had one or more new intakes during the specified period. *Add_intk3* reports any new intakes within three months (90 days) of the initial intake; *add_intk_acc3* counts only accepted intakes (FAR or investigative), excluding Screened Out or Risk Only intakes. We also created separate variables (*add_intk_out3*, *add_intk_FAR3*, *add_intk_invst3*, and *add_intk_risk3*), which correspond to the number of Screened Out, FAR or FAR-eligible, Investigative non-FAR-eligible, and Risk Only intakes.

Costs

To address the cost research question, we generated a continuous non-negative variable representing the dollar expenditure on purchased goods and services for each family made by CA. The file of services included a row for each service provided to each family, and also included the cost of each service. We aggregated—for the appropriate time periods of 3, 6, 12, or 24 months after intake—the total cost of services provided to each family.

The *serv_cost3* variable represents the cost of services provided at any date between the intake date and three months after the intake date. Using the same process, we created outcome variables for other time periods: *serv_cost6* for cost incurred within 6 months and analogous variables for 12 and 24 months.

Selecting Data for Matching of FAR Families to Non-FAR Comparison Families

To perform matching of FAR families to FAR-eligible investigative families, we requested information on potentially relevant family characteristics. Our data request focused on any family characteristic that could change the effect of FAR on our measured outcomes, including variables related to prior economic assistance, prior involvement with CA, criminal histories of family members, mental health and medical histories, and many other similar factors. CA provided over 300 covariates in the cohort files, with many representing variations on the same variable (e.g., covering all time periods versus only the previous year) and the total amount of some activity versus binary indicators of any amount.

Because binary versions of these variables would reduce variability and therefore decrease the precision of estimates, we utilized continuous versions when available. For financial assistance variables, we selected a single variable representing total assistance from all Washington State sources. The final list of covariates used in propensity score matching and as control variables in multiple regression is as follows:

Final List of Matching Variables	
County Urbanization	Level of urbanization of the county of the FAR office in which the family receives services based on US Department of Agriculture designations
Criminal Involvement	Number of family members with any criminal involvement prior to FAR intake (any time prior)
Criminal Severity	The severity of the most severe criminal offense of any family member prior to FAR intake (any time prior)
Disability (DD) Eligibility	Number of family members eligible for disability benefits
Domestic Violence History	Number of family members with a domestic violence charge prior to FAR intake (any time prior)
Emergency Room Use	Total number of family members using emergency room care (number of visits) prior to FAR intake (any time period)
First CA Encounter	(Yes/No) Indicates whether this is the first CA encounter for any family member
Homelessness History	Total number of household members experiencing homelessness prior to FAR intake (any time period)
Injury History	Total number of injuries reported to any family member prior to FAR intake (any time period)

Final List of Matching Variables	
Intake Type	Type of Intake (Neglect/Abandonment, Physical Abuse, Sexual Abuse/Exploitation)
Juvenile Justice History	Total number of prior adjudications for all juvenile family members prior to FAR intake (any time prior)
Medical/Medicaid Eligibility	Number of months eligible for medical assistance (maximum for family member) prior to FAR intake
Mental Health History	Total number of family members with mental health diagnosis prior to FAR intake (any time prior)
Mental Health History (Severity)	Most severe mental health diagnosis across family members prior to FAR intake (any time prior)
Number of Children	Count of the number of children living with the family at time of FAR intake
Prior AOD Treatment	Total number of times family member(s) (any) were treated for alcohol or other drug issues prior to FAR intake (any time prior)
Prior Economic Assistance	Sum of family's total economic assistance received prior to FAR intake (any time prior)
Race/Ethnicity (Youngest Child)	Race/ethnicity of youngest child in the family, as recorded in FamLink
Risk Scores	Abuse and neglect scores derived from SDM Risk Assessment
Tribal Affiliation	CA flag indicating an Indian Child Welfare case
Youngest Child's Age	Age of the youngest child in the family at the time of intake

We generated the variable representing the number of children using data provided in the *far_persons* data set, which contains information on every person related to an intake. Using this data set, we calculated the age of every person involved in any intake and excluded those individuals 18 and over. After eliminating any observations with the same *ID_CASE* (family ID) and *ID_PRSN* (person ID), we summed the number of children by *ID_CASE*. We added this TriWest-generated variable into each cohort file. In early versions of the data, many families did not have any children listed in the *far_persons* file. This problem was reduced significantly in later updates of the data.

There are two risk score variables, *abuse* and *neglect*, each based on risk scores completed through the Structured Decision-Making (SDM) risk assessment. The cohort data set includes the date the SDM risk information was entered into FamLink. A comparison of intake dates to SDM dates demonstrated that SDM information was entered on average approximately 45 days after intake rather than at the beginning of the case. Because the entered information may be results of the intervention, instead of pre-existing family characteristics before the intervention, we did not use the CA-generated neglect or abuse risk scores as matching or control variables. We instead separated those components of the risk scores that were based on unchanging characteristics (such as number of prior intakes) and developed our own risk and abuse scores. Many observations contained missing values.

"Youngest Child's Age" was drawn from the cohort variable *ageintk_yngst*, which represents the age of the youngest family member. In the first two cohorts of data we initially received, this variable contained many missing values or had values that were contradictory (e.g., negative ages or adult ages). We replaced problematic values by using values from the previously mentioned *far_persons* data set. More recent transfers of data have substantially fewer missing or errant values after replacement from the *far_persons* data.

Imputation

While deletion of observations or variables with missing values is the most common practice in econometric analysis, the current state of the art is to impute missing values when the variables in question are statistically important and contain more than a trivial number of missing values. Early analysis of the *ageintk_yngst* variable, which in our first two data sets contained thousands of missing values, convinced us that excluding observations with missing values had the potential for biasing our measurements of the effect of FAR on removal rates. The variables representing abuse or neglect risk scores continue to include several thousand missing values, as does race of the youngest child. While imputing missing values adds significant complexity to the analysis, concerns about bias convinced us of the value of imputation.

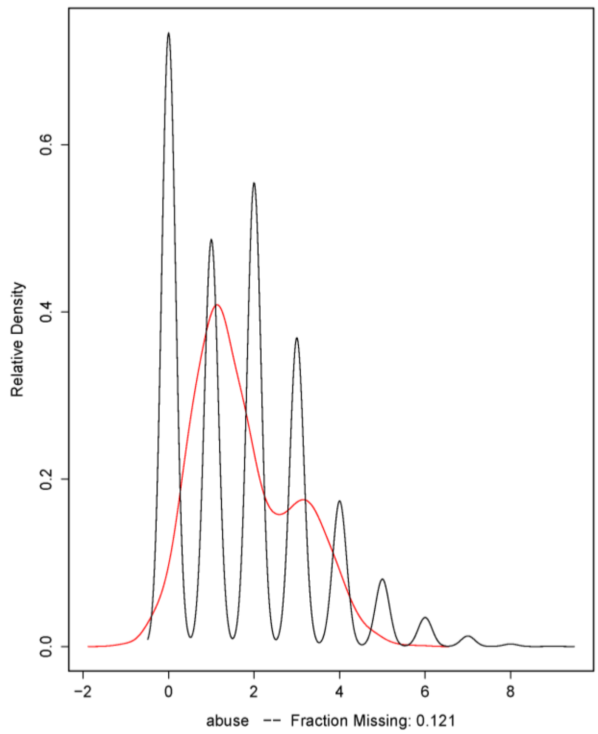
The software program we use, Amelia, performs multiple imputations. It uses non-missing data to estimate the likely distributions of the missing data and then creates multiple data sets that are identical for the non-missing data but contain unique values for the missing data, each randomly drawn from the estimated distributions. We used five imputed data sets (the default number). When using these multiple data sets for outcome analysis, we analyze each data set separately, then combine results across the data sets in a manner that accounts for the additional uncertainty of missing data. The process for sample averages and regression coefficients is to simply average the results. For standard errors, the combined standard error includes the average standard error plus a measurement of the variability in the sample means or regression coefficients. We refer to these combination procedures as "Rubin's Rules," referring to Donald Rubin and colleagues, who, in work going back to the 1970s, demonstrated that under a broad range of conditions yielding missing data, multiple imputation yields results that are unbiased and efficient.[1]

The imputation process generated data for four continuous variables with missing values plus the categorical variable race of the youngest child. The following figures provide information for

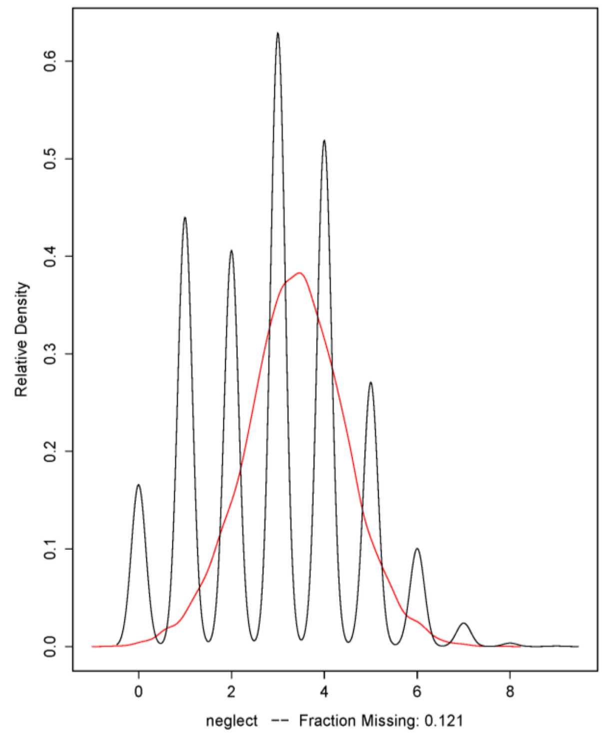
[1] See King et al. for an explanation of the advantages and limitations of multiple imputation: King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. *American Political Science Review*, *95*(1), 49–69.

the continuous variables on the distribution of the imputed (red) and non-missing (black) data, and the proportion of data that were imputed.

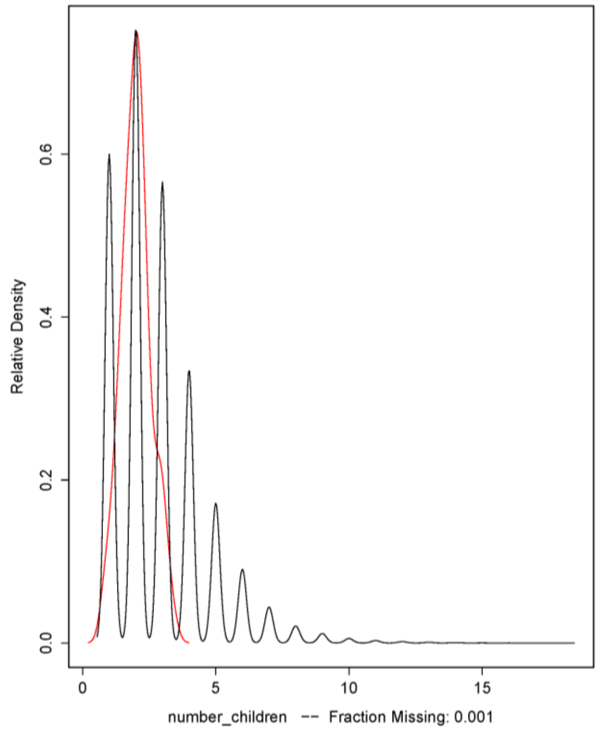
Observed and Imputed Values of Abuse



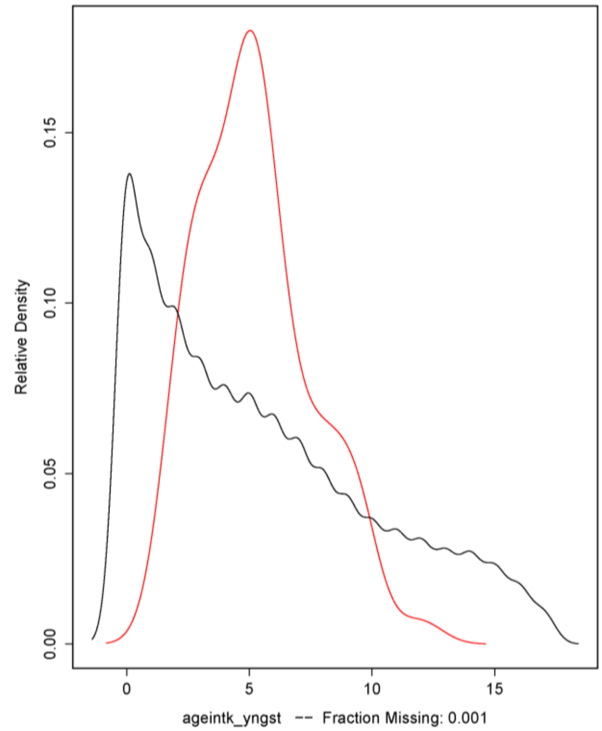
Observed and Imputed Values of Neglect



Observed and Imputed Values of Number_Children



Observed and Imputed Values of ageintk_yngst



Note that in the data used for this report, only a small fraction of the variables *number of children* and *age of the youngest child* were missing. Missing values for abuse and neglect risk scores were still very high, and imputing missing values is likely to have reduced bias in our outcome analysis. This process was not, however, without flaws. As we describe in the outcome analysis section, performing statistical tests on five slightly different data sets, then combining the results, adds substantial complexity. For some simple output, such as counts of events, policy makers will want straightforward answers, while we are compelled to give answers that are averages across the five data sets. These complexities make reporting of results more complicated and reduce credibility when presenting results to audiences without backgrounds in econometrics and statistics.

Propensity Score Matching

As per the approved evaluation plan, outcome analysis was performed after propensity score matching of treatment (FAR) to comparison (FAR-eligible investigative) observations. Each of the five imputed data sets per cohort was run separately through the R program, MatchIt. This program works by running a logistic regression with the binary *farcase* indicator (indicating the family was actual FAR) as the dependent variable and the matching covariates as the independent variables. No outcome variables were used. Based on this logistic regression, the fitted value of each observation was calculated. Since the dependent variable is binary, the fitted value for each observation is the probability that that observation was in the treatment group (*farcase* = 1) based on the values of the covariates. This probability is the observation's propensity score.

MatchIt then ranks all treatment and comparison observations by propensity score and, starting at the top, matches each treatment observation to the comparison observation with the closest propensity score. We performed a one-to-one match: each FAR family was matched to one FAR-eligible investigative family. Unmatched investigative families were discarded from the data set.

Standard Bias Confirmation

As part of our evaluation of the reliability of our propensity score matching results and methodology—and to assess balance of the covariate values in the treatment and comparison samples—we calculated standard bias estimates on all 36 variables used in propensity score matching. These estimates were calculated for both pre- and post-matching samples. We use standard bias estimates to address three topics of concern:

- Are the unmatched treatment and comparison pools sufficiently similar in covariate values to allow direct comparison of outcomes? Did we need to control for differences through propensity score matching?
- Does matching reduce differences in the average covariate values of the treatment and comparison groups?
- Is the post-matching comparison group sufficiently comparable to allow direct comparison of outcomes to the treatment group, or should we use multiple regression in the outcome analysis to further control for differences in covariate values?

Standard bias estimates compare treatment group (FAR) covariate distributions to comparison group (FAR-eligible investigative families) covariate distributions. The covariates analyzed are the same as those used to match the treatment and comparison families. The standard bias estimates are calculated by differencing the treatment and comparison covariate means and dividing that difference by the standard deviation of the difference.[2] To properly assess balance—whether the respective treatment and comparison groups are adequately similar and suitable for comparison—we used a threshold value of 0.1. Standard bias estimates of less than 0.1 were taken to indicate good balance (similarity) in a matching variable's distribution between the treatment and comparison groups. A standard bias threshold of 0.1 indicates that

[2] Standard bias continuous variable = $(|mean_t - mean_c|) / sqrt(((sd_t)^2 + (sd_c)^2) / 2)$
Standard bias binary variable = $(|prop_t - prop_c|) / sqrt(prop_t * (1-prop_t) + prop_c * (1-prop_c))$

any difference in means between the two groups is less than 10% of one standard deviation. We have selected this threshold as the most reasonably conservative one, as Harder, Stuart, and Anthony note that a cutoff of 0.25 is an acceptable standard but that 0.1 is preferable in propensity score matching.[3]

While it is also possible to calculate percentage change in standard bias from pre- to post-matching, very small adjustments to standard bias in real terms may be disproportionately represented by reporting percent change. As such, we do not consider percent change reliable for gauging change in balance that results from matching. Instead, we present the pre- and post-match standard bias estimates for all 36 matching variables and assess results by comparing estimates to the 0.1 threshold.

It is important to note that standard bias only compares the sample means or proportions of the treatment and comparison groups. Other features of each distribution, such as variance or skewness, are not directly assessed with standard bias. In the case of binary variables, which are the majority of the covariates we use, comparison of sample proportions is sufficient since the sample proportion provides all useful information about the covariate's distribution. Most of the remaining covariates are discrete counts (e.g., number of children) involving small numbers. Because these variables will not have extreme values (e.g., 10,000 children in a family), large differences in variance or skewness are unlikely.

We calculated standard bias estimates both by cohort (resulting in seven sets of estimates—one for each cohort) and for our data in aggregate (across all cohorts). We deemed this two-part approach prudent because propensity score matching occurs cohort-by-cohort, but the outcome analysis is based on data for all cohorts pooled into a single data set. While we calculated cohort level standard bias estimates, the results were similar to those of the data in aggregate, and we are only reporting the aggregate results.

In the table on the following page, we sort the 36 covariates into four color-coded categories. Those in dark green represent covariates with standard bias estimates that improved as the result of matching but were below the 0.1 threshold (i.e., well matched) before matching. Those in light green also represent improvements in balance from matching, but the improvement was from above the threshold (poor matches) to below the threshold (good matches).

---

[3] Harder, V. S., Stuart, E. A., & Anthony, J. C. (2010). Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychological Methods*, *15*(3), 6. https://doi.org/10.1037/a0019623

Covariates in gray represent cases in which matching worsened balance, but the post-matching standard bias remained below the threshold value. Cases in orange represent cases in which pre- and post-matching standard bias estimates are above the threshold.

- **Issue 1.** Are the treatment and comparison pools sufficiently similar in covariate values to allow direct comparison of outcomes? Or do we need to control for differences through propensity score matching?

This issue is addressed by examining those covariates with standard bias estimates that are above the threshold before any matching was conducted. The covariates highlighted in orange (3) and light green (13) make up 16 of 36 (44%) of the covariates. Because such a large percent of the covariates are unbalanced, some form of matching or other control for incomparability of treatment and comparisons groups is warranted.

- **Issue 2.** Does matching reduce differences in the average covariate values of the treatment and comparison groups?

Matching did reduce standard bias from above to below the threshold for the 11 covariates highlighted in light green. But it failed to reduce bias below the threshold for the 3 covariates highlighted in orange. Propensity score matching was therefore helpful at reducing potential bias but insufficient to eliminate it for all covariates. For 5 covariates highlighted in gray, matching increased standard bias but not to levels above the threshold.

- **Issue 3.** Is the post-matching comparison group sufficiently comparable to allow direct comparison of outcomes to the treatment group, or should we use multiple regression in the outcome analysis to further control for differences in covariate values?

The three covariates highlighted in orange retain standard bias estimates above the threshold of 0.1, but all are below the less conservative standard of 0.25. The three variables are *Any_ESA_sum*, which represents the sum of all forms of economic assistance; *county_urban5*, representing child welfare offices in rural locations; and *criminvolve_N*, representing the number of family members with criminal involvement.

Because these three variables are likely to be correlated with outcomes, there remains the potential for omitted variable bias if analysis of outcomes does not further control for these. In our outcome analysis, we report simple comparisons of outcomes for the treatment and matched comparison groups, and we perform analysis with multiple regression to further control for group differences in covariate values. Results from both forms of analysis are generally similar, and in the few examples in which the analyses gave differing results, the multiple regression results are less likely to contain bias.

Standard Bias Estimates for Propensity Matching Variables, Pre- and Post-Matching		
Matching Variable	All Data	
	Pre	Post
Risk Score Abuse	0.014	0.013
Youngest Child's Age	0.001	0.017
Prior Economic Assistance	0.158	0.107
Prior AOD Treatment	0.11	0.05
Asian Pacific Islander	0.007	0.042
Black	0.027	0.027
County Urbanization 2	0.2	0.057
County Urbanization 3	0.486	0.02
County Urbanization 4	0.039	0.009
County Urbanization 5	0.174	0.19
Criminal Involvement	0.34	0.134
Criminal Severity	0.255	0.085
Disability (DD) Eligibility	0.062	0.039
Domestic Violence History	0.072	0.037
Emergency Room Use	0.119	0.07
Hispanic	0.052	0.006
Homelessness History	0.088	0.04
First CA Encounter	0.101	0.059
Injury History	0.135	0.077
Intake Type Abuse	0.015	0.008
Intake Type Sexual Abuse	0.045	0.045
Juvenile Justice History	0.208	0.081
Medical/Medicaid Eligibility	0.085	0.08
Mental Health History	0.199	0.099
Mental Health Severity 1	0.079	0.001
Mental Health Severity 2	0.002	0.004
Mental Health Severity 3	0.148	0.051
Mental Health Severity 4	0.029	0.011
Mental Health Severity 5	0.112	0.041
Multiracial Asian	0.01	0.006
Multiracial Black	0.023	0.003
Multiracial Native American	0	0.004
Native American	0.072	0.014
Risk Score Neglect	0.125	0.026

Standard Bias Estimates for Propensity Matching Variables, Pre- and Post-Matching		
Matching Variable	All Data	
	Pre	Post
Number of Children	0.017	0.054
Tribal Affiliation	0.096	0.007

The following table presents means for the same variables, January through June 2017.

Standard Bias Estimates for Propensity Matching Variables, Means			
Matching Variable	Mean FAR	Mean All Investigative	Mean Matched Comparison
Propensity Score	0.476	0.231	0.357
Abuse	1.620	1.695	1.598
ageintk_yngst	5.985	5.556	6.147
Any_ESA_sum	211.972	211.765	202.916
aodtx_sum	10.052	8.210	9.268
county_urban2	0.200	0.023	0.048
county_urban3	0.128	0.283	0.152
county_urban4	0.048	0.016	0.030
county_urban5	0.036	0.005	0.012
criminvolve_N	1.124	0.744	1.002
crimsevr_max	1.424	1.111	1.403
ddelg_N	0.256	0.244	0.258
domviol_sum	0.224	0.194	0.286
ERuse_sum	17.500	17.590	16.114
Hmls_max	11.560	13.459	11.884
incep_allfm	0.180	0.200	0.194
injury_sum	20.788	18.168	18.073
intk_abuse	0.348	0.366	0.351
intk_sex_abuse	0.000	0.000	0.000
Juvcrim_sum	1.520	1.104	1.353
medelg_max	79.500	85.065	77.828
MI_Broad_Npers	1.748	1.564	1.668
MIrxsvr1	0.032	0.041	0.038
MIrxsvr2	0.064	0.074	0.051
MIrxsvr3	0.352	0.269	0.348
MIrxsvr4	0.004	0.009	0.004
MIrxsvr5	0.184	0.170	0.163
Neglect	2.627	3.092	2.621

number_children	2.569	2.784	2.488
Native.American	0.031	0.040	0.034
Asian.Pacific.Islander	0.050	0.066	0.065
Black	0.097	0.125	0.111
Hispanic	0.090	0.158	0.082
Multiracial.Native.American	0.030	0.044	0.030
Multiracial.Black	0.062	0.053	0.064
Multiracial.Asian	0.016	0.012	0.017
Watribe	0.032	0.041	0.034

Outcome Analysis

With propensity score matching complete, the seven cohorts were combined, with binary cohort indicators added and used as additional covariates. We analyzed the effect of FAR on the probability of a removal, additional intakes (i.e., re-referrals), and costs.

Our analysis approach was to perform a simple difference in means test (T test) or proportions test (chi-squared test) between the FAR treatment and matched comparison groups. This testing was then followed by a more sophisticated regression-based test. We use regression-based tests for two reasons: (1) a continuing potential lack of comparability of covariate distributions of the FAR treatment and matched comparison group (see above) and (2) extreme skewness in the distribution of some outcome variables, leading to potential bias in T and chi-squared tests. The simple tests we conducted were limited by the distributions of the outcome variables; the continuous cost variables were very highly skewed. Given outcome data that were dominated by zeros and were highly skewed, T-tests have the potential to yield biased estimates of the effect of FAR.

In conducting regression-based tests, we utilized the same set of covariates used in propensity score matching. For purposes of reporting, we described magnitudes of the effect of FAR by comparing the expected values for each observation (both FAR and comparison families) when the binary FAR indicator was set to 1.0 (FAR treatment) versus set to 0.0 (comparison group). We differenced these expected values to measure the effect of FAR for each family in the data. We then averaged across all families this difference in expected values. Additional details on this process are reported below. We measured statistical significance of FAR by T tests on the FAR treatment indicator coefficient in the regression output. We do not report the estimated linear equations or the statistical significance of other covariates.

For purposes of comparison, we report the simple and regression-based results on the same table.

Removals

We calculated removal rates using the previously described outcome variables, *removal3, removal6, removal12, and removal24*. These binary variables indicated whether a family had one or more removals within the time period indicated (e.g., three months for *removal3*).

Because both the dependent variables and the treatment variable *farcase* are binary, we conducted a simple test of a difference in proportions of families with a removal with a chi-squared test. The sample proportion of FAR and comparison families with a removal is reported in the following table. This proportion figure represents the average of the sample proportions of the five imputed data sets. The magnitude of the difference between FAR and comparison families is reported under “Magnitude of Effect: Chi-Squared.” The P value of the difference between FAR and comparison families generated via chi-squared test is reported under “P Value: Chi-Squared.”

Using this simple test of a difference in proportions, we found that (1) FAR families had lower removal rates for all four periods and (2) the effect of FAR on removals was negative. Differences were statistically significant at conventional significance levels at 3, 6, and 12 months. The P value of the chi-squared test for 24-month differences was 0.336 and is not statistically significant.

Removal Outcome Analysis Without Separate Cohort Treatment								
Time Range	Sample Proportions		Logistic Regression Expected Value		Magnitude of FAR Effect		P Value	
	FAR	Comparison	FAR	Comparison	Chi-Squared	Logistic Regression	Chi-Squared	Logistic Regression
3 months	0.029	0.041	0.029	0.041	-0.012	-0.012	0.000	0.000
6 months	0.043	0.055	0.043	0.055	-0.012	-0.012	0.001	0.001
12 months	0.060	0.073	0.060	0.073	-0.013	-0.013	0.004	0.002
24 months	0.087	0.093	0.086	0.094	-0.006	-0.007	0.336	0.207

We also evaluated the effect of FAR on removals with logistic regression, using the same five matched data sets and the matching variables (plus the binary cohort indicator) as covariates. Instead of reporting sample proportions of FAR and comparison families, we report the expected values for all families had they entered the FAR pathway rather than the investigative pathway. Our procedure for calculating expected value was to first use logistic regression to estimate a linear equation of the effect of all covariates, including the binary FAR treatment variable, on the probability of a removal. For each family in the data set, both FAR and comparison, we set the FAR variable to the value 1.0 (indicating the family was FAR) and calculated the probability that the family would have a removal using the estimated equation and each family’s covariate values. We then set each family’s FAR variable to 0.0 and re-recalculated the probability of a removal.

As a result of this process, for each family we had two expected values; each represents the probability of a removal given the family’s covariate values, with one case in which the family was assumed as FAR and the second as not-FAR (i.e., comparison). The average of these expected values is reported under “Logistic Regression Expected Value.” The difference in expected values averaged over all families is reported under “Magnitude of FAR Effect: Logistic Regression.”

The P value of the binary FAR treatment variable is reported in the last column of the previous table under "P Value: Logistic Regression." This P value is taken directly from the logistic regression output and may be interpreted as the probability of observing an effect size of the reported magnitude when the true effect size is zero. In the case of removals within three months of intake, the estimated effect of FAR is to reduce the probability of a removal by 0.012—or 1.2 percentage points. The probability of drawing a random sample from a population with a sample effect of 1.2 percentage points, when the population level effect is really zero, is the P value, reported as zero (0.0000436 unrounded).

Generally, logistic regression yielded similar results to chi-squared tests. FAR families had lower removal rates, and this effect was statistically significant at 3, 6, and 12 months. The magnitude of the reduction at 12 months (-0.013) measured as a percentage of the comparison rate (0.073) was approximately 18%, a modest but promising reduction.

In the logistic regression results reported in the previous table, we included binary variables indicating the cohort of the FAR and comparison families. This allowed removal rates to vary over time, but this cohort-to-cohort variation was measured as if it were the same for both FAR and comparison families within each cohort. As an example, removal rates for both FAR families and comparison families may have been lower in the second cohort than in the first cohort; including a binary cohort variable would capture that cohort-to-cohort variation. We are also interested in estimating whether the effect of FAR on removals varies by cohort (i.e., does the difference between FAR and comparison families vary by cohort?). It is plausible that because of improved training and experience with the FAR program, the outcome results of FAR would improve over time.

By interacting the binary FAR variable with the binary cohort variables and using logistic regression, we were able to measure separate effects of FAR by cohort. With this approach, both the average removal rate and the effect of FAR could vary by cohort. The following table (next page) reports the expected value of FAR and comparison families, measured as previously described, and the magnitude of the effect of FAR. We also report the regression coefficients on each interaction variable (the cohort variable multiplied by the FAR treatment variable). Because the equation used for estimating separate effects of FAR by cohort differs from the equation we use when we assume FAR has the same effect in each cohort, the expected values reported below differ slightly from the previous table.

Removals with Separate Cohort Treatment						
Time Range	Logistic Regression Expected Value		Magnitude of Aggregate FAR Effect	Cohort	Cohort FAR Effects	P Value
	FAR	Comparison				
3 months	0.029	0.041	-0.012	**Cohort 1**	-0.503	0.169
				Cohort 2	-0.246	0.110
				Cohort 3	-0.369	0.052
				Cohort 4	-0.899	0.002
				Cohort 5	-0.098	0.730
				Cohort 6	-0.223	0.534
				Cohort 7	-1.601	0.043
6 months	0.044	0.056	-0.012	**Cohort 1**	-0.488	0.086
				Cohort 2	-0.154	0.247
				Cohort 3	-0.247	0.122
				Cohort 4	-0.669	0.006
				Cohort 5	-0.151	0.541
				Cohort 6	-0.195	0.538
12 months	0.061	0.075	-0.013	**Cohort 1**	-0.431	0.096
				Cohort 2	-0.082	0.465
				Cohort 3	-0.263	0.056
				Cohort 4	-0.593	0.005
				Cohort 5	-0.040	0.839
24 months	0.091	0.098	-0.007	**Cohort 1**	-0.398	0.105
				Cohort 2	0.012	0.906
				Cohort 3	-0.134	0.273

The numbers reported under the "Cohort FAR Effects" column are the regression coefficients on the terms interacting the binary FAR indicator with the binary cohort variable. Negative coefficients represent a reduction in the probability of a removal for FAR versus comparison families in that cohort; positive coefficients indicate an increase in probability. The magnitude of each coefficient does not have any clear meaning. For example, for three-month removals and Cohort 1, the regression coefficient is -0.503. This means that, as compared to comparison families in Cohort 1, FAR families had a reduction in the log odds ratio of 0.503. Since log odds is a non-linear transformation, we cannot interpret this to imply a 0.503 reduction in the probability of a removal. These coefficients may be used comparatively. For three-month outcomes, Cohort 4 had the largest reduction in the probability of a removal.

In reviewing the cohort-specific regression coefficients, the following patterns stand out. First, Cohorts 1, 3, 4, and 7 had the largest reductions in the probability of a removal. Cohorts 2, 5, and 6 had smaller magnitudes. Next, Cohorts 3 and 4 had the lowest P values, indicating the increased likelihood that the identified reduction in removals attributable to FAR was not a result of sampling error.

The high P values (i.e., low probability that the measured effect was real) for Cohorts 1 and 7 were likely caused by their small sample sizes. Cohort 1 had very few FAR families. Cohort 7 had very few FAR-eligible investigative families. Each limitation reduced the number of matched pairs, which increases the difficulty of measuring the effect of FAR with precision.

Discounting Cohorts 1 and 7, the pattern for removals over time appears to be better results on removals for FAR during the middle of the intervention—during Cohorts 3 and 4. The effect of FAR on removals appears to be driven by these cohorts, with smaller measured effects that are not statistically significant in earlier and later cohorts.

Additional Intakes

As described previously, we measured additional intakes or re-referrals using multiple outcome variables. These measures included any additional intakes, accepted intakes, and intakes broken out by type (Screened Out, FAR or FAR-eligible, non-FAR-eligible investigative, Risk Only). We used the binary form of outcome or dependent variables, indicating that one or more intakes of the specified type occurred within the time period. In earlier versions of the analysis, we measured the dependent variables as counts of the numbers of intakes for each family. Analysis results were similar whether using binary or count versions of the variables, and for ease of interpretation, we selected the binary form.

In analyzing additional intakes, we again first used chi-squared tests to determine if any differences are statistically significant. The next table is organized analogously to the removals table. The sample proportion of FAR families with an accepted intake within three months of the initial cohort intake was 0.126, while comparison families had a sample proportion of 0.113. These FAR families were more likely to have subsequent accepted intakes within three months of the cohort intake. The magnitude of the effect of FAR is an increase in the proportion of families with an accepted intake of 0.012. When using chi-squared tests, this difference is statistically significant with a P value of 0.024.

While an examination of accepted intakes suggests that FAR increased the probability of future intakes (an outcome inconsistent with the program goals), examination of subsequent FAR-eligible versus non-FAR-eligible investigative intakes provides a more nuanced understanding of the impact of FAR. In particular, FAR appears to increase the probability of subsequent FAR (or FAR-eligible) intakes. But FAR reduces the probability of non-FAR-eligible investigative intakes. For example, using 12-month results and sample proportions, the probability of an accepted FAR or FAR-eligible intake is 0.072 higher for FAR families than for comparison families. But the probability of an accepted non-FAR-eligible investigative intake is 0.016 lower for FAR families. Since FAR eligibility is driven in large part by the seriousness of the allegation, these results suggest that FAR increases the probability of future intakes, but reduces the seriousness of the allegations.

This pattern (i.e., higher probability of FAR-eligible intakes but lower probability of non-FAR-eligible investigative intakes) is consistent across the 3-, 6-, 12-, and 24-month time periods. The accepted intake results are statistically significant during all four periods, as is the higher probability of FAR and FAR-eligible intakes. The statistical significance of lower non-FAR-eligible investigative intakes loses statistical significance at 24 months.

Outcome Analysis with Additional Intakes as Binary Variables

Time Period/ Intake Type	Sample Proportion		Expected Value (Logistic Regression)		Magnitude Effect of FAR		P Value	
	FAR	Comparison	FAR	Comparison	Chi-Squared	Logistic Regression	Chi-Squared	Logistic Regression
3 months								
Accepted Intakes	0.126	0.113	0.125	0.114	0.012	0.011	0.024	0.057
Screened Out	0.193	0.180	0.190	0.183	0.012	0.007	0.054	0.284
FAR Intakes	0.095	0.066	0.094	0.067	0.029	0.028	0.000	0.000
Investigative Intakes	0.039	0.056	0.038	0.056	-0.017	-0.018	0.000	0.000
Risk Only Intakes	0.007	0.007	0.007	0.007	0.000	0.000	0.684	0.767
6 months								
Accepted Intakes	0.194	0.166	0.192	0.167	0.028	0.025	0.000	0.000
Screened Out	0.263	0.242	0.259	0.246	0.021	0.013	0.004	0.093
FAR Intakes	0.145	0.099	0.144	0.100	0.046	0.044	0.000	0.000
Investigative Intakes	0.069	0.086	0.068	0.086	-0.017	-0.019	0.000	0.000
Risk Only Intakes	0.012	0.015	0.012	0.014	-0.002	-0.002	0.249	0.263
12 months								
Accepted Intakes	0.275	0.226	0.272	0.228	0.049	0.044	0.000	0.000
Screened Out	0.345	0.311	0.339	0.316	0.034	0.023	0.000	0.004
FAR Intakes	0.209	0.136	0.206	0.137	0.072	0.069	0.000	0.000
Investigative Intakes	0.110	0.126	0.109	0.127	-0.016	-0.018	0.006	0.001
Risk Only Intakes	0.024	0.027	0.023	0.028	-0.004	-0.004	0.166	0.102

Outcome Analysis with Additional Intakes as Binary Variables								
Time Period/ Intake Type	Sample Proportion		Expected Value (Logistic Regression)		Magnitude Effect of FAR		P Value	
	FAR	Comparison	FAR	Comparison	Chi-Squared	Logistic Regression	Chi-Squared	Logistic Regression
24 months								
Accepted Intakes	0.371	0.283	0.368	0.286	0.088	0.081	0.000	0.000
Screened Out	0.450	0.396	0.444	0.402	0.054	0.042	0.000	0.000
FAR Intakes	0.288	0.171	0.285	0.173	0.116	0.112	0.000	0.000
Investigative Intakes	0.165	0.172	0.163	0.173	-0.007	-0.010	0.377	0.200
Risk Only Intakes	0.047	0.047	0.046	0.048	0.000	-0.002	0.770	0.706

In addition to chi-squared tests, we also used logistic regression to measure the effect of FAR on the probability of one or more subsequent intakes. As in the case of removals, regression-based tests allowed us to control for confounding covariates that have the potential of biasing the results of a simple chi-squared test. The logistic regression results were very similar to those of chi-squared. We measured expected value using the same approach described with removal analysis.

Service Costs: Cost of FAR Versus Matched Comparison Families

We measured the costs of CA-provided goods and services. These did not include the cost of Children's Administration staff time and were not divided into costs used to assist families (e.g., the purchase of concrete goods or family therapy versus the cost of providing foster care).

Since cost is a non-negative continuous variable, we used T tests for a simple measurement of difference in means between FAR and comparison families. The average cost for each group is reported on the following page in the "Service Costs Analysis without Separate Cohort Treatment" table, under "T-Test Sample % or Average." The difference in these sample averages is reported under "Magnitude of Effect: T-Test." Based on sample averages, FAR families had higher costs at 3 months ($35 per family) but lower costs at 6, 12, and 24 months (-$103, -$346, and -$717 respectively). The 6- and 12-month results are statistically significant, while 3- and 24-month results are not.

The underlying distribution of costs per family is highly skewed and zero-dominated. The five-number summary (plus mean) of the 12-month costs is displayed in the table on the following page:

Cost Per Family Distribution					
Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
$0	$0	$0	$950	$0	$206,300

Note that the third quartile value remains zero; this distribution is dominated by no expenditures on families. The mean of $950 is not representative of "typical" spending on families (which is $0). The small number of families with very large expenditures drives the average expenditure of $950.

Given this distribution, T tests have the potential for biasing estimates of effect size and statistical significance. We selected a "hurdle" model that allowed the same underlying variables (our matching variables, plus the FAR indicator variable) to separately estimate the probability of any expenditures (the first hurdle) and the size of those expenditures (the second hurdle). This approach allowed FAR to have differential effects; it could increase the probability of any expenditures while reducing the magnitude of expenditures for those families with positive values.

We used a Probit model to estimate the first hurdle. For the second hurdle, we used log of expenditures as the dependent variable and the same matching variables as the independent variables. Wooldridge[4] provided the expected values for this econometric model, which we used to calculate expected values for each family.

	Service Costs Analysis without Separate Cohort Treatment										
	T Test Sample % or Average		Hurdle Expected Value		Magnitude of Effect				P Value		
	FAR	Comparison	FAR	Comparison	T-Test	Hurdle 1	Hurdle 2	Combined	T-Test	Hurdle 1	Hurdle 2
3 months	$238	$203	$326	232	$35	0.543	-0.425	$95	0.060	0.000	0.000
6 months	$403	$506	$591	636	-$103	0.508	-0.737	-$45	0.009	0.000	0.000
12 months	$831	$1177	$1,120	$1,515	-$346	0.472	-0.870	-$394	0.000	0.000	0.000
24 months	$2,168	$2,885	$2,640	$4,127	-$717	0.464	-0.946	-$1487	0.011	0.000	0.000

The coefficients reported under "Magnitude of Effect: Hurdle 1, Hurdle 2" are the regression coefficients on the FAR treatment indicator variable. A positive coefficient indicates FAR increased the probability of any expenses (Hurdle 1) or the amount of expenses for families with positive amounts (Hurdle 2). For all time periods in this analysis, FAR increased the

[4] Wooldridge, J. M. (2001). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass: MIT Press. 537.

probability of incurring expenditures while reducing the amount of the expenditures. The net effect, the expected value of expenses after controlling for all of the covariates, is reported under "Combined." FAR increased CA expenditures for 3 months after intake. By 6 months, families that received FAR, after controlling for covariates, had expenses that were on average $45 lower than what they would have been if these same families had received investigations. These results are statistically significant for all time periods.

Using the same hurdle model, we used interaction terms to measure separate cohort effects. The pattern for each cohort was the same: FAR increased the probability of positive expenditures while decreasing the amount of expenditures for families with positive expenditures. Most, but not all, of the cohort hurdles were statistically significant, and the 12-month magnitude of FAR was a reduction of $416 per family. The difference between this amount and the $394 predicted previously reflects a more complex underlying linear model, which allows the effect of FAR to vary by cohort.

Service Cost Analysis with Separate Cohort Treatment									
Hurdle Regression Expected Value		Proportion of Positive Values		Magnitude of Effect		Hurdle 1		Hurdle 2	
FAR	Comparison	FAR	Comparison		Cohort	Cohort Effect	P Value	Cohort Effect	P Value
3 months									
$331	$233	0.213	0.096	$98	**Cohort 1**	0.691	0.000	-0.409	0.027
					Cohort 2	0.564	0.000	-0.489	0.000
					Cohort 3	0.422	0.000	-0.445	0.002
					Cohort 4	0.396	0.000	-0.461	0.008
					Cohort 5	0.605	0.000	-0.461	0.156
					Cohort 6	0.706	0.000	-0.261	0.329
					Cohort 7	0.753	0.000	-0.521	0.187
6 months									
$598	$646	0.239	0.120	$-48	**Cohort 1**	0.638	0.000	-0.578	0.011
					Cohort 2	0.540	0.000	-0.800	0.000
					Cohort 3	0.398	0.000	-0.837	0.000
					Cohort 4	0.378	0.000	-0.718	0.000
					Cohort 5	0.571	0.000	-0.718	0.002
					Cohort 6	0.644	0.000	-0.646	0.020
12 months									
$1,148	$1,564	0.266	0.146	$-416	**Cohort 1**	0.630	0.000	-0.658	0.006
					Cohort 2	0.508	0.000	-0.961	0.000
					Cohort 3	0.355	0.000	-0.861	0.000
					Cohort 4	0.355	0.000	-0.873	0.000
					Cohort 5	0.585	0.000	-0.873	0.000

Service Cost Analysis with Separate Cohort Treatment									
Hurdle Regression Expected Value		Proportion of Positive Values		Magnitude of Effect		Hurdle 1		Hurdle 2	
FAR	Comparison	FAR	Comparison		Cohort	Cohort Effect	P Value	Cohort Effect	P Value
24 months									
$2,743	$4,267	0.307	0.177	$-1,524	**Cohort 1**	0.594	0.000	-0.728	0.001
					Cohort 2	0.519	0.000	-0.955	0.000
					Cohort 3	0.344	0.000	-0.971	0.000

This pattern of increased proportion of FAR families with positive expenditures (but lower expenditures for those with any expenditure) was consistent with the underlying FAR model. The focus of the intervention was to provide services and supports to families in order to address underlying problems instead of waiting until an episode of child abuse or neglect requires more expensive interventions, such as removals and placements in foster care. Consistent with this understanding of FAR, removals were lower at 3, 6, and 12 months (see above). We currently are unable to confirm that lower costs for FAR are driven by lower removal rates but hope to perform this additional analysis for the final evaluation report.

The econometric analysis of the FAR cohort data yields results that are broadly consistent with the anticipated outcomes and with outcomes in other states. For periods 3, 6, and 12 months after intake, the probability of a removal is lower for FAR families than for matched comparison families. The number of subsequent intakes increases for all time periods, but the number of non-FAR investigative intakes are generally reduced.

In other states, the effect of FAR on costs has been inconsistent. The complexity of the underlying relationship may be a source of this inconsistency. FAR does increase the probability of some costs, but for families with positive costs, FAR reduces the magnitude of those costs.