Appendix C

EVALUATION OF *FAMILY-TO-FAMILY* USING PROPENSITY SCORE ANALYSIS: A COMPARISON STUDY

Shenyang Guo, Ph.D.
School of Social Work
University of North Carolina at Chapel Hill

The evaluation of the anchor-site phase of *Family to Family* basically employs a covariance control approach, that is, it runs multivariate analysis using covariates as control variables plus an indicator of Family to Family (F2F) key elements versus No F2F key elements to measure the average treatment effect of *Family to Family*. Researchers have found that the conventional covariance control approach has numerous flaws and should be replaced by more rigorous methods in drawing causal inference. For instance, Sobel (1996) criticized the common practice in sociology that uses a dummy variable (i.e., treatment versus nontreatment) to evaluate the treatment effect in a regression model (or a regression-type model such as a path analysis or structural equation modeling) using survey data.

The primary problems of covariance control approach discussed in the literature may be summarized as follows: (1) the dummy treatment variable is specified by these models as exogenous, but in fact it is not, and determinants of incidental truncation or sample selection should be explicitly modeled first, and selection effects should be taken into consideration when estimating causal impacts on outcomes (Heckman, 1978, 1979); (2) the strongly ignorable treatment assignment assumption (i.e., conditional upon covariates, the treatment assignment is independent from outcomes under both treatment and control conditions) is prone to violation in observational studies; under such condition, the presence of the endogeneity problem leads to a biased and inconsistent estimation of the regression coefficient (Berk, 2004; Imbens, 2004; Rosenbaum & Rubin, 1983); and (3) covariance control does not automatically correct for nonignorable treatment assignment (Guo & Fraser, 2009).

Over the past 30 years, methods of program evaluation have undergone significant change as researchers have recognized the need to develop more efficient approaches for assessing treatment effects from studies based on observational data and for evaluations based on quasi-experimental designs. Propensity score analysis is one of such latest advances that have proven useful for program evaluation when randomized clinical trials are infeasible or unethical. Because of data

limitations and resource constraints, the evaluation team did not apply this approach in analyses presented in the final report. While the team was confident about the assumptions that guided the decision to employ a covariance control approach, questions remain as to the extent to which propensity score analysis might alter study findings and lead to different conclusions? To answer this question, the evaluation team invited us to analyze a subset of the *Family to Family* data (i.e., children receiving protective services from Maricopa County, Arizona) and conduct an in-depth comparison study. studying response, this report compares findings generated by the covariance control approach with those generated by 7 propensity score approaches.

The data for this analysis pertain to 18,185 children who entered into protective services in Maricopa County of Arizona between 2002 and 2007. The Family to Family treatment variable was created from data extracted from a database used to capture information about key elements of Family to Family that were apparent at the time of F2F key elements meetings held to decide either (1) whether a child should be removed from the home of their birth family due maltreatment; or (2) whether a child in out-of-home care should be moved to another placement setting. Based on the implementation schedule followed in Maricopa County, those who entered into services between 2005 and 2007 were considered to have some level of exposure to key elements of Family to Family practice. In contrast, children who were served between 2002 and 2004 were considered to be a pre-implementation comparison group.

The study analyzed four outcome variables: (1) hazard rate of reunification within 12 months; (2) hazard rate of reunification/exit to relatives within 12 months; (3) likelihood of reentry into foster care 6 months after reunification; and (4) likelihood of being initially placed in family settings. Because the first two outcomes are time-to-event data with random and right-hand censoring, we employed Cox proportional hazards model in the analysis. We employed logistic regression to analyze the remaining two outcomes.

The independent variables employed in these models, in addition to the treatment dichotomous variable, include: gender, age at first entry, race (i.e., three indicator variables of “African American”, “Hispanic”, and “Other”, with “White” as a reference group), type of initial placement (i.e., five indicator variables of “relatives licensed”, “relatives unlicensed and own home”, “group home”, “emergency shelter”, and “other”, with “foster home” as a reference group), and indicator variables of entry cohorts. Since the entry cohorts are a linear function of the treatment variable “F2F key

elements", we created four indicator variables to code cohorts 2003, 2004, 2005, and 2006, and used cohorts 2002 and 2007 as a reference group. Since the cohort variables serve as a control and it is the F2F key elements indicator that indicates average treatment effect, the coding scheme combining entry cohorts 2002 and 2007 does not create methodological problem.

We analyzed the four outcomes on each of the following eight samples to achieve the analytic objective of model comparison.

1. ***Original*** – the original sample without using propensity score modeling. This is the sample our entire evaluation employed. For the current comparison, we intended to see to which extent findings from the original sample would differ from the remaining seven samples using propensity score modeling.

2. ***Greedy 1*** – the sample was created by a propensity score greedy matching known as nearest-neighbor within caliper matching (Rosenbaum and Rubin, 1983), where we used a binary logistic regression to estimate *logit* propensity scores, and used a quarter of one standard deviation of the estimated logit propensity score as a caliper size (i.e., the caliper size e=.25s_p=.0591).

3. ***Greedy 2*** – the sample was created by the same method as that for Greedy 1 (i.e., a nearest-neighbor within caliper matching), but we used a latest development in the field, known as the generalized boosted regression (GBR), to estimate *logit* propensity scores (Friedman, Hastie, & Tibshirani, 2000). The primary advantage of GBR is that the method is insensitive to misspecification of wrong functional forms of predictor variables in a logistic regression. We used a quarter of one standard deviation of the estimated logit propensity score as a caliper size (i.e., the caliper size e=.25s_p=.0111).

4. ***Greedy 3*** – the sample was created by the same method as that for Greedy 2 (i.e., a nearest-neighbor within caliper matching, where the logit propensity scores were estimated by GBR), but we used a narrowed specification of the caliper size (i.e., the caliper size e=.005, or a half of the caliper size for Greedy 2) to test sensitivity of findings to caliper.

5. ***PSW 1 (ATE)*** – the sample was created by propensity score weighting (McCaffrey, Ridgeway, & Morral, 2004), where we used the same logit propensity scores as those for Greedy 1 (i.e., the propensity scores were estimated by a logistic regression), and employed weights for estimating average treatment effect (ATE) in the weighted analysis.

6. ***PSW 1 (ATT)*** – the sample was created by the same procedure as PSW 1 (ATE), but we employed weights for estimating average treatment effect for the treated (ATT) in the weighted analysis.

7. ***PSW 2 (ATE)*** – the sample was created by the same procedure as PSW 1 (ATE), but we used the propensity scores created by GBR (i.e., the same scores as Greedy 2)

and employed weights for estimating average treatment effect (ATE) in the weighted analysis.

8. ***PSW 2 (ATT)*** – the sample was created by the same procedure as PSW 1 (ATE), but we used the propensity scores created by GBR (i.e., the same scores as Greedy 2) and employed weights for estimating average treatment effect for the treated (ATT) in the weighted analysis.

Details of the analytic methods and a detailed version of the findings are presented below. Exhibit C.1 summarizes the main findings of the comparison study. As Exhibit C.1 shows, results clearly reveal a high convergence of study findings among the 8 samples, in terms of directions of treatment effect and patterns of statistical significance tests. Specifically, all 8 samples indicate that: children exposed to F2F key elements had a higher hazard rate of achieving reunification within one year than children not exposed to F2F key elements ($p<.001$); children exposed to F2F key elements had a higher hazard rate of reunification or exit to relatives within one year than children not exposed to F2F key elements ($p<.001$); and children exposed to F2F key elements had higher odds of being initially placed in family settings than children not exposed to F2F key elements ($p<.05$, or $p<.01$, or $p<.001$). All 8 samples studying the likelihood of reentry into foster care within 6 months do not show significant difference between children exposed to F2F key elements and those not exposed to F2F key elements.

It is important to note a limitation embedded in the propensity score analysis: unlike randomized clinical trials, the propensity score approach cannot correct for selection bias generated by unmeasured variables (Rubin, 1997). Thus, it is likely that if we controlled for additional and more important variables affecting selections, results using the 7 propensity-score samples might be different, and be different in a way that distinguishes them from the original sample.

In summary, within the data and design constraints that prevail in the current evaluation of *Family to Family*, analyses using the original sample with a conventional covariance control approach are valid, and do not show findings that are different from those generated by propensity score analysis.

Exhibit C.1: Comparison of Estimated Treatment Effects among Models

Treatment Effect and Outcome	Original		Greedy 1		Greedy 2		Greedy 3		PSW1(ATE)		PSW1(ATT)		PSW2(ATE)		PSW2(ATT)	
Hazard ratio of F2F key elements on reunification within one year	1.328	***	1.312	***	1.298	***	1.302	***	1.322	***	1.329	***	1.327	***	1.329	***
Hazard ratio of F2F key elements on exit to relatives within one year	1.454	***	1.450	***	1.412	***	1.401	***	1.458	***	1.473	***	1.456	***	1.464	***
Odds ratio of F2F key elements on reentry within 6 months	1.155		1.133		1.128		1.079		1.169		1.184		1.158		1.169	
Odds ratio of F2F key elements on initial placement in family settings	1.651	***	1.149	*	1.789	***	1.678	***	1.257	***	1.233	**	1.575	***	1.421	***

* p<.05, ** p<.01, *** p<.001, two-tailed

Methods

To draw valid causal inference, this study applies the Neyman-Rubin counterfactual framework (Neyman, 1923; Morgan & Winship, 2007; Rubin, 1974, 2006) as a conceptual model to guide the data analysis. Under this setting, a counterfactual is a *potential* outcome, or the state of affairs that would have happened in the absence of the cause (Shadish, Cook, & Campbell, 2002); and a *counterfactual framework* emphasizes that individuals selected into either treatment or nontreatment groups have potential outcomes in both states: that is, the one in which they are observed and the one in which they are not observed. The Neyman-Rubin framework offers a practical way to evaluate the counterfactuals. Working with event data from a sample that represents the population of interest, the standard estimator for the average treatment effect is seen as the difference between two estimated medians from the sample data as:

$$\hat{t} = Median\ \ (\hat{T}_1 \mid w = 1) - Median\ \ (\hat{T}_0 \mid w = 0),$$

where $\hat{T}_1$ is the event time under the treated condition, $\hat{T}_0$ is the event time under the control condition, and w is a binary variable indicating treatment receipt (i.e., w=1, treatment; and w=0, control).

To estimate the average treatment effect for the treated, the estimator becomes:

$$\hat{t} = [(Median\ \ \hat{T}_1 - Median\ \hat{T}_0) \mid w = 1]$$

The counterfactual framework underscores the importance of balancing data, carefully seeking the potential outcome (equivalently choosing a most comparable comparison group) for a valid causal inference, and estimating appropriate treatment effects using appropriate methods suiExhibit to research questions.

This study employs the following methods to balance data to draw a valid causal inference: (1) a propensity score greedy matching (i.e., the nearest neighbor within caliper matching) followed by a survival analysis or a logistic regression analysis; (2) a refined approach called generalized boosted regression (GBR) to estimate the propensity scores, and use the same greedy matching to create matched samples to conduct a follow-up survival analysis or a logistic regression analysis; and (3) propensity score weighting applied to a survival analysis or a logistic regression. Key features of these propensity score approaches are described below.

The Definition of Propensity Scores

With complete data, Rosenbaum and Rubin (1983) defined the propensity score for participant *i* (*i*=1, ..., *N*) as the conditional probability of assignment to a particular treatment (W_i=1) versus nontreatment (W_i=0) given a vector of observed covariates, $\boldsymbol{x_i}$:

$$e(\boldsymbol{x}_i) = pr(W_i=1 \mid \boldsymbol{X_i} = \boldsymbol{x_i})$$

In this study, we employed a logistic regression or GBR to estimate conditional probability of receiving treatment (i.e., exposed to F2F key elements) for each study child. Following Rosenbaum and Rubin (1985), we defined the propensity score for analyses using greedy matching as a logit score. Denoting the predicted probability of receiving treatment as $\hat{e}(x)$, the logit propensity scores for samples generated by Greedy 1, Greedy 2, and Greedy 3 are defined as:

$\hat{q}(x) = \log[(1-\hat{e}(x))/\hat{e}(x)]$. According to Rosenbaum and Rubin (1985), a logit propensity score $\hat{q}(x)$ is better than a score using the predicted probability, because the distribution of the logit propensity score approximates to normal. In the following text, unless specifically pointed out, we will use the term "logit propensity scores" interchangeably with "propensity scores".

Propensity Score Greedy Matching. The method of propensity score greedy matching (Rosenbaum & Rubin, 1983, 1985) involves the following steps. First, it uses the binary logistic regression or GBR to estimate a propensity score of receiving treatment (i.e., exposed to F2F key elements). The propensity score is a balancing score representing a vector of covariates or the so-called "conditioning variables". The advantage of the propensity score matching is its reduction of dimensions. The conditioning variables the study aims to match may include many covariates. The propensity score approach reduces all this dimensionality to a one-dimensional score. Doing so, it eases the burden of finding matches within the study sample. The logistic regression predicting propensity scores employed all covariates available for this study, or the same set of independent variables used in the analysis of the original sample.

Second, it matches the treated participants to controls on the estimated propensity scores to make the estimate of counterfactuals (i.e., outcome values of the comparison group) more valid. This study employs the nearest neighbor within a caliper matching (Rosenbaum & Rubin, 1985). The method selects a control participant *j* as a match for treated participant *i*, if and only if the absolute distance of propensity scores between the two participants (i.e., the difference between propensity scores P_i and P_j) meets the following condition:

$$\| P_i - P_j \| < e,$$

where *e* is a prespecified tolerance for matching, or a caliper. Rosenbaum and Rubin (1985) suggest using a caliper size of a quarter of a standard deviation of the sample estimated propensity scores (i.e., $\varepsilon \leq .25\sigma_P$, where σ_P denotes standard deviation of the estimated propensity scores of the sample). We followed this recommendation in Greedy 1 and Greedy 2, but employed a narrowed caliper size in Greedy 3 to test whether or not findings may change.

Finally, based on the matched sample, the study conducts the Cox proportional hazards model or the binary logistic regression model to study outcome differences between treated and comparison participants.

Generalized Boosted Regression. Estimating propensity scores plays a central role in propensity score analysis. In addition to using the binary logistic regression, this study employs the generalized boosted regression (GBR, also known as generalized boosted modeling), to fulfill the task. One of the problems with the binary logistic regression is specifying an unknown functional form for each predictor. If specifying functional forms can be avoided, then the search of a best model involves fewer subjective decisions and, therefore, may lead to a more accurate prediction of treatment probability.

GBR is a general, automated, data-adaptive algorithm that fits several models by way of a regression tree, and then merges the predictions produced by each model. As such, GBR can be used with a large number of pretreatment covariates to fit a nonlinear surface and predict treatment assignment. GBR is one of the latest prediction methods which have been rapidly made popular in the machine learning community as well as mainstream statistics research (Guo & Fraser, 2009). From a statistical perspective, the breakthrough in applying boosting to logistic regression and exponential family models was made by Friedman, Hastie, and Tibshirani (2000). These authors showed that an exponential loss function used in a machine-learning algorithm *AdaBoost* was closely related to the Bernoulli likelihood. Thus, GBR offers a good alternative to logistic regression in which specifying the unknown functional form of predictor variables is avoided.

The key feature and advantage of GBR is that the analyst does not need to specify functional forms of the predictor variables. GBR employs regression trees. As McCaffrey et al. (2004) pointed out, trees handle continuous, nominal, ordinal, and missing independent variables, and they capture

nonlinear and interaction effects. A useful property of trees is that they are invariant to one-to-one transformations of the independent variables. Thus, "whether we use age, log(age), or age^2 as a participant's attribute, we get exactly the same propensity score adjustments" (McCaffrey et al., 2004, p. 408). This property explains why uncertainty about a correct functional form for each predictor variable is no longer an issue when GBR is used.

This study employed the Stata ***boost*** program developed by Schonlau (2007) to estimate the GBR propensity scores.

Propensity Score Weighting Analysis. Propensity score weighting treats estimated propensity score $\hat{e}(x)$ [i.e., the propensity score using predicted probability, rather than the logit score $\hat{q}(x)$] as a sampling weight, and incorporate the weights in multivariate analysis (McCaffrey, et al., 2004). It aims to reweight treated and control participants to make them representative of the population of interest –a procedure similar to estimators for stratified sampling. The crucial element of analysis is the development of weights based on the estimated propensity scores. Different types of weights are used, depending on whether an average treatment effect or the average treatment effect for the treated is desired.

1. For estimating the average treatment effect (ATE), we define weights as follows:

$$w(W,x)=\frac{W}{\hat{e}(x)}+\frac{1-W}{1-\hat{e}(x)}$$

By this definition, when W=1 (i.e., a treated participant), the weight becomes $w(W,x)=\frac{1}{\hat{e}(x)}$;

and when W=0 (i.e., a control), the weight becomes $w(W,x)=\frac{1}{1-\hat{e}(x)}$.

2. For estimating the average treatment effect for the treated (ATT), we define weights as follows:

$$w(W,x)=W+(1-W).\frac{\hat{e}(x)}{1-\hat{e}(x)} \quad .$$

By this definition, when W=1 (i.e., a treated participant), the weight becomes $w(W,x)=1$; and

when W=0 (i.e., a control), the weight becomes $w(W,x)=\frac{\hat{e}(x)}{1-\hat{e}(x)}$.

We applied the above weights to the Cox regression and logistic regression to discern treatment effects. The analyses were implemented by Stata version 10.

Findings

Exhibits C.2 and C.3 show the histograms and box-plots of the estimated propensity scores by treatment status. As the figures indicate, the two groups share a sizeable common-support region, and the greedy matching may not reduce the size of a matched sample to a large degree.

Exhibit C.4 presents descriptive statistics of the original sample (i.e., the sample before matching) and matched samples using the propensity score greedy matching. Results show that Greedy 1 excluded 2,027 participants (11.1%), Greedy 2 excluded 2,361 participants (13.0%), and Greedy 3 excluded 2,495 participants (13.7%) from the outcome analyses. Because the matched samples range from 15,824 to 16,158, the outcome analyses are valid and each matched sample in general represents the original one.

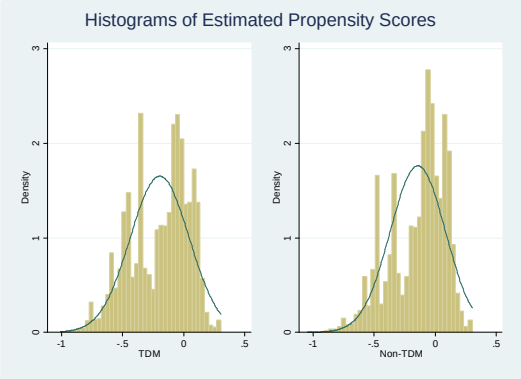
As Exhibit C.5 shows, matching improves covariate balances. Race and type of initial placement are statistically significant at a .001 level before matching. After matching, most of these differences are no longer significant at the same level. There remain differences that are statistically significant. Much of these differences are attributable to large sample size. In short, matching has improved covariate balances, and the follow-up outcome analyses based on the matched samples should be less sensitive to selection bias caused by these covariates.

Exhibits C.5 to C.8 present the final results of the four outcome analyses using the 8 samples. Results clearly show a high convergence of study findings among the 8 samples, in terms of directions of treatment effect and patterns of statistical significance tests. Specifically, all 8 samples reveal that: children exposed to F2F key elements had a higher hazard rate of achieving reunification within one year than children not exposed to F2F key elements ($p<.001$); children exposed to F2F key elements had a higher hazard rate of reunification or exit to relatives within one year than children not exposed to F2F key elements ($p<.001$); and children exposed to F2F key elements had higher odds of being initially placed in family settings than children not exposed to F2F key elements ($p<.05$, or $p<.01$, or $p<.001$). All 8 models studying the likelihood of reentry into foster care within 6 months do not show significant difference of reentry between children exposed to F2F key elements and those not exposed to F2F key elements.

Rubin (1997) summarized three limitations of propensity score matching, one of which is that propensity score matching cannot control for unobserved selection bias. This limitation applies to the current study. Given the data constraints, we only used gender, age, race, and type of initial placement in the propensity score analysis. The covariates controlled by this study may not encompass all or most important factors affecting sample selections. Thus, we must interpret the findings with caution, and condition them on the limitation that the method fails to correct for hidden selections. It is likely that using additional covariates in matching or weighting, we may obtain different results and conclusions.

In summary, within the data and design constraints that prevail in the current evaluation of *Family to Family*, analyses using the original sample with a conventional covariance control approach are valid, and do not show findings that are different from those generated by propensity score analysis.

Exhibit C.2: Distribution of the Propensity Scores Estimated by the Logistic Regression



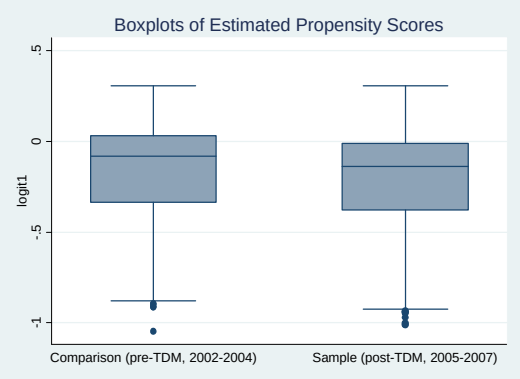
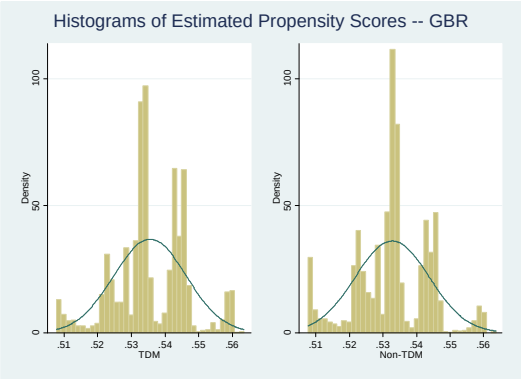


Exhibit C.3: Distribution of the Propensity Scores Estimated by the Generalized Boosted Regression



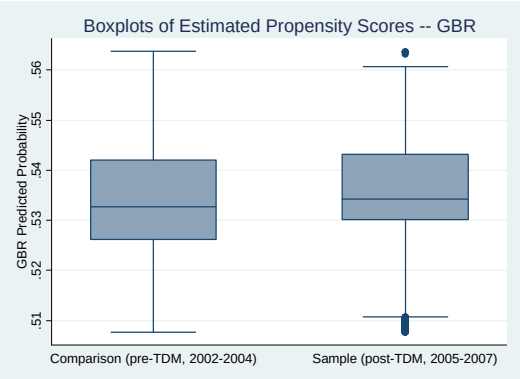


Exhibit C.4: Checking Covariate Balance Before and After Matching (Entries are % of F2F key elements Participants in the Indicated Group, or Mean[SD])

Matching Variable	Before Matching	After Greedy 1 Matching	After Greedy 2 Matching	After Greedy 3 Matching
N - Non-TDM Participants	8,329	8,079	7,912	7,845
N - TDM Participants	9,856	8,079	7,912	7,845
N of Participants Lost after Matching(%)		2,027(11.1%)	2,361(13.0%)	2,495(13.7%)
Gender				
Female	54.6%	49.7%	51.0%*	51.4%**
Male	53.9%	50.3%	49.0%*	48.7%**
Age at entry (in years)				
TDM Participants	7.35(5.98)	7.59(5.94)	7.38(5.82)*	7.30(5.90)**
Non-TDM Participants	7.48(5.89)	7.51(5.91)	7.57(5.95)*	7.58(5.96)**
Race				
White	51.3%***	49.1%	49.9%	50.4%
African American	57.7%***	50.0%	49.2%	49.8%
Hispanic	55.1%***	50.4%	49.4%	48.7%
Other	62.1%***	45.2%	51.4%	48.8%
Type of Initial Placement				
Foster home	61.4%***	46.1%*	46.6%	53.5%**
Relative licensed	64.3%***	45.8%*	51.5%*	48.8%
Relative unlicensed & own home	51.7%***	50.8%	51.5%*	51.9%**
Group home	48.8%***	51.4%	52.6%**	51.6%
Emergency shelter	51.6%***	51.2%	49.1%	48.8%*
Other	59.1%***	49.3%	8.7%***	24.7%***

*** p<.001, ** p<.01, * p<.05, chi-square test or independent-sample t test two-tailed

Exhibit C.5: Estimated Hazard Ratios from Cox Regression for Reunification within One Year: Comparisons of the Uncorrected Model Using Original Sample with Corrected Propensity Score Models

Covariate	Original		Greedy 1		Greedy 2		Greedy 3		PSW 1 (ATE)		PSW 1 (ATT)		PSW 2 (ATE)		PSW 2 (ATT)	
	HR	p	HR	p	HR	p	HR	p	HR	p	HR	p	HR	p	HR	p
Gender (Female)																
Male	0.996	0.850	0.989	0.573	0.994	0.780	0.996	0.829	0.996	0.733	0.997	0.815	0.996	0.747	0.997	0.795
Age at entry (in years)	1.024	0.000	1.024	0.000	1.025	0.000	1.025	0.000	1.024	0.000	1.024	0.000	1.024	0.000	1.024	0.000
Race (White)																
African American	1.037	0.266	1.038	0.283	1.049	0.171	1.042	0.233	1.046	0.040	1.050	0.029	1.038	0.087	1.042	0.059
Hispanic	1.023	0.262	1.017	0.414	1.016	0.466	1.019	0.378	1.022	0.086	1.022	0.096	1.022	0.080	1.022	0.083
Other	1.264	0.000	1.175	0.020	1.279	0.000	1.254	0.001	1.252	0.000	1.257	0.000	1.257	0.000	1.256	0.000
Type of Initial Placement (Foster)																
Relative licensed	1.160	0.004	1.170	0.011	1.169	0.012	1.145	0.026	1.151	0.001	1.151	0.001	1.152	0.001	1.151	0.001
Relative unlicensed & own home	0.832	0.000	0.820	0.000	0.823	0.000	0.816	0.000	0.829	0.000	0.829	0.000	0.828	0.000	0.829	0.000
Group home	0.834	0.000	0.823	0.000	0.821	0.000	0.815	0.000	0.831	0.000	0.832	0.000	0.831	0.000	0.832	0.000
Emergency shelter	1.007	0.815	0.991	0.787	0.996	0.902	0.990	0.747	1.004	0.829	1.008	0.668	1.006	0.738	1.008	0.691
Other	0.953	0.284	0.917	0.081	0.850	0.007	0.870	0.014	0.931	0.013	0.934	0.020	0.939	0.028	0.937	0.024
Cohort (2002 & 2007)																
2003	0.866	0.000	0.865	0.000	0.870	0.000	0.871	0.000	0.871	0.000	0.875	0.000	0.868	0.000	0.871	0.000
2004	0.841	0.000	0.841	0.000	0.840	0.000	0.841	0.000	0.841	0.000	0.841	0.000	0.842	0.000	0.842	0.000
2005	0.629	0.000	0.638	0.000	0.649	0.000	0.651	0.000	0.635	0.000	0.632	0.000	0.631	0.000	0.630	0.000
2006	0.668	0.000	0.678	0.000	0.675	0.000	0.678	0.000	0.673	0.000	0.670	0.000	0.669	0.000	0.669	0.000
Family-to-Family (No TDM)																
TDM	1.328	0.000	1.312	0.000	1.298	0.000	1.302	0.000	1.322	0.000	1.329	0.000	1.327	0.000	1.329	0.000
N	18,178		16,152		15,817		15,683		18,178		18,178		18,178		18,178	
% of subjects lost due to PSA			11.1%		13.0%		13.7%		0.0%		0.0%		0.0%		0.0%	

Note: HR - Hazard Ratio; p - p value from two-tailed test; reference group is shown in parenthesis for categorical variable.

Exhibit C.6: Estimated Hazard Ratios from Cox Regression for Exit to Relatives within One Year: Comparisons of the Uncorrected Model Using Original Sample with Corrected Propensity Score Models

Covariate	Original		Greedy 1		Greedy 2		Greedy 3		PSW 1 (ATE)		PSW 1 (ATT)		PSW 2 (ATE)		PSW 2 (ATT)	
	HR	p	HR	p	HR	p	HR	p	HR	p	HR	p	HR	p	HR	p
Gender (Female)																
Male	0.994	0.698	0.990	0.549	0.994	0.729	0.990	0.572	0.995	0.684	0.996	0.777	0.994	0.646	0.995	0.707
Age at entry (in years)	1.021	0.000	1.022	0.000	1.023	0.000	1.023	0.000	1.022	0.000	1.021	0.000	1.022	0.000	1.021	0.000
Race (White)																
African American	1.041	0.137	1.037	0.216	1.022	0.460	1.016	0.590	1.032	0.152	1.030	0.176	1.036	0.110	1.034	0.134
Hispanic	0.977	0.174	0.966	0.055	0.967	0.070	0.965	0.053	0.975	0.050	0.976	0.063	0.975	0.054	0.976	0.059
Other	1.284	0.000	1.231	0.000	1.217	0.000	1.218	0.000	1.254	0.000	1.255	0.000	1.274	0.000	1.267	0.000
Type of Initial Placement (Foster)																
Relative licensed	1.085	0.068	1.091	0.108	1.105	0.060	1.115	0.036	1.072	0.077	1.068	0.095	1.078	0.053	1.075	0.063
Relative unlicensed & own home	0.669	0.000	0.675	0.000	0.664	0.000	0.667	0.000	0.673	0.000	0.671	0.000	0.669	0.000	0.670	0.000
Group home	0.923	0.010	0.921	0.015	0.909	0.004	0.912	0.007	0.929	0.002	0.932	0.004	0.923	0.001	0.927	0.002
Emergency shelter	1.182	0.000	1.189	0.000	1.181	0.000	1.185	0.000	1.194	0.000	1.201	0.000	1.184	0.000	1.192	0.000
Other	0.850	0.000	0.869	0.001	0.774	0.000	0.808	0.000	0.844	0.000	0.845	0.000	0.843	0.000	0.844	0.000
Cohort (2002 & 2007)																
2003	0.900	0.001	0.895	0.000	0.897	0.001	0.896	0.001	0.904	0.000	0.908	0.000	0.902	0.000	0.904	0.000
2004	0.908	0.002	0.906	0.001	0.901	0.001	0.900	0.001	0.910	0.000	0.911	0.000	0.909	0.000	0.910	0.000
2005	0.651	0.000	0.648	0.000	0.668	0.000	0.677	0.000	0.652	0.000	0.649	0.000	0.651	0.000	0.650	0.000
2006	0.697	0.000	0.701	0.000	0.707	0.000	0.713	0.000	0.700	0.000	0.697	0.000	0.698	0.000	0.697	0.000
Family-to-Family (No TDM)																
TDM	1.454	0.000	1.450	0.000	1.412	0.000	1.401	0.000	1.458	0.000	1.473	0.000	1.456	0.000	1.464	0.000
N	18,178		16,152		15,817		15,683		18,178		18,178		18,178		18,178	
% of subjects lost due to PSA			11.1%		13.0%		13.7%		0.0%		0.0%		0.0%		0.0%	

Note: HR - Hazard Ratio; p - p value from two-tailed test; reference group is shown in parenthesis for categorical variable.

Exhibit C.7: Estimated Odds Ratios from Logistic Regression for Reentry within Six Months: Comparisons of the Uncorrected Model Using Original Sample with Corrected Propensity Score Models

Covariate	Original		Greedy 1		Greedy 2		Greedy 3		PSW 1 (ATE)		PSW 1 (ATT)		PSW 2 (ATE)		PSW 2 (ATT)	
	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p	OR	p
Gender (Female)																
Male	0.988	0.862	0.999	0.993	0.957	0.575	0.992	0.922	0.998	0.982	0.987	0.856	0.996	0.957	0.989	0.883
Age at entry (in years)	0.972	0.000	0.975	0.001	0.968	0.000	0.967	0.000	0.972	0.000	0.971	0.000	0.972	0.000	0.972	0.000
Race (White)																
African American	0.844	0.142	0.785	0.062	0.786	0.067	0.804	0.093	0.813	0.075	0.808	0.067	0.829	0.104	0.822	0.090
Hispanic	0.782	0.002	0.801	0.009	0.759	0.001	0.754	0.001	0.772	0.001	0.766	0.001	0.773	0.001	0.770	0.001
Other	0.578	0.012	0.369	0.002	0.443	0.006	0.450	0.007	0.557	0.008	0.555	0.007	0.572	0.010	0.566	0.009
Type of Initial Placement (Foster)																
Relative licensed	1.132	0.442	1.123	0.571	0.979	0.918	1.011	0.957	1.005	0.975	1.006	0.973	1.081	0.630	1.053	0.749
Relative unlicensed & own home	0.588	0.001	0.631	0.006	0.657	0.009	0.624	0.004	0.595	0.001	0.602	0.001	0.577	0.000	0.589	0.001
Group home	1.051	0.716	1.101	0.532	1.161	0.312	1.154	0.339	1.067	0.636	1.085	0.549	1.058	0.679	1.069	0.624
Emergency shelter	1.075	0.487	1.181	0.169	1.171	0.159	1.145	0.237	1.100	0.358	1.114	0.298	1.086	0.419	1.098	0.364
Other	0.633	0.022	0.618	0.036	0.832	0.556	0.922	0.754	0.664	0.042	0.662	0.040	0.647	0.029	0.653	0.033
Cohort (2002 & 2007)																
2003	0.891	0.405	0.899	0.453	0.863	0.300	0.846	0.243	0.881	0.361	0.866	0.309	0.890	0.400	0.877	0.343
2004	0.766	0.058	0.728	0.028	0.746	0.043	0.731	0.031	0.759	0.051	0.748	0.044	0.766	0.057	0.755	0.047
2005	0.755	0.013	0.729	0.012	0.719	0.009	0.783	0.055	0.752	0.012	0.754	0.013	0.757	0.014	0.754	0.013
2006	0.740	0.007	0.706	0.006	0.660	0.001	0.708	0.008	0.727	0.005	0.741	0.008	0.740	0.008	0.741	0.008
Family-to-Family (No TDM)																
TDM	1.155	0.257	1.133	0.350	1.128	0.367	1.079	0.573	1.169	0.218	1.184	0.187	1.158	0.246	1.169	0.217
N	6,657		5,809		5,696		5,634		6,657		6,657		6,657		6,657	
% of subjects lost due to PSA			12.7%		14.4%		15.4%		0.0%		0.0%		0.0%		0.0%	

Note: OR - Odds Ratio; p - p value from two-tailed test; reference group is shown in parenthesis for categorical variable.

Exhibit C.8: Estimated Odds Ratios from Logistic Regression for Initially Placed in Family Setting: Comparisons of the Uncorrected Model Using Original Sample with Corrected Propensity Score Models

	Original		Greedy 1		Greedy 2		Greedy 3		PSW 1 (ATE)		PSW 1 (ATT)		PSW 2 (ATE)		PSW 2 (ATT)	
Covariate	**OR**	**p**	**OR**	**p**	**OR**	**p**	**OR**	**p**	**OR**	**p**	**OR**	**p**	**OR**	**p**	**OR**	**p**
Gender (Female)																
Male	0.842	0.000	0.870	0.000	0.890	0.001	0.896	0.002	0.843	0.000	0.844	0.000	0.845	0.000	0.846	0.000
Age at entry (in years)	0.850	0.000	0.853	0.000	0.849	0.000	0.847	0.000	0.850	0.000	0.846	0.000	0.850	0.000	0.848	0.000
Race (White)																
African American	0.821	0.000	0.646	0.000	0.723	0.000	0.874	0.025	0.838	0.002	0.845	0.004	0.820	0.000	0.824	0.001
Hispanic	1.094	0.012	1.000	0.993	0.967	0.374	0.969	0.417	1.097	0.010	1.097	0.011	1.096	0.011	1.093	0.014
Other	0.663	0.000	0.508	0.000	0.543	0.000	0.471	0.000	0.688	0.000	0.680	0.000	0.668	0.000	0.662	0.000
Cohort (2002 & 2007)																
2003	1.174	0.011	1.183	0.008	1.218	0.002	1.206	0.004	1.156	0.021	1.145	0.035	1.178	0.009	1.170	0.013
2004	0.992	0.892	1.009	0.888	1.033	0.608	1.024	0.718	0.958	0.492	0.933	0.278	0.992	0.900	0.969	0.615
2005	0.624	0.000	0.651	0.000	0.510	0.000	0.538	0.000	0.626	0.000	0.621	0.000	0.620	0.000	0.622	0.000
2006	0.925	0.162	0.906	0.112	0.802	0.000	0.827	0.003	0.919	0.127	0.924	0.152	0.922	0.141	0.924	0.154
Family-to-Family (No TDM)																
TDM	1.651	0.000	1.149	0.036	1.789	0.000	1.678	0.000	1.257	0.000	1.233	0.001	1.575	0.000	1.421	0.000
N	18,185		16,158		15,824		15,690		18,185		18,185		18,185		18,185	
% of subjects lost due to PSA			11.1%		13.0%		13.7%		0.0%		0.0%		0.0%		0.0%	

Note: OR - Odds Ratio; p - p value from two-tailed test; reference group is shown in parenthesis for categorical variable.

References

Berk, R. A. (2004). *Regression analysis: A constructive critique*. Thousand Oaks, CA: Sage.

Friedman, J., Hastie, T., & Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting (with discussion). *Annals of Statistics, 28*, 337-374.

Guo, S. & Fraser, W.M. (2009). *Propensity score analysis: Statistical methods and Applications*. Thousand Oaks, CA: Sage Publications.

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equations system. *Econometrica, 46*, 931-960.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*, 153-161.

Imbens, G.W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics, 86*, 4-29.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods, 9*, 403-425.

Morgan, S.L, & Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.

Neyman, J. S. (1923). Statistical problems in agricultural experiments. *Journal of the Royal Statistical Society Series B, 2:* 107-180.

Rosenbaum, P. R. (2002). *Observational studies* (2nd ed.). New York: Springer.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*, 33-38.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*, 688-701.

Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127*, 757-763.

Rubin, D. B. (2006). *Matched sampling for causal effects*. New York: Cambridge University Press.

Schonlau, M. (2007). *Boost module for Stata*. Retrieved August 22, 2008, from http://www.stata-journal.com/software/sj5-3

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Sobel, M. E. (1996). An introduction to causal inference. *Sociological Methods & Research, 24*, 353-379.

Appendix D

Anchored Scales Used to Assess Progress in Implementing Core Strategies

Site: _____ Team Decision Making				
1	**2**	**3**	**4**	**5**
Has not yet begun to hold TDM meetings. Has created a TDM strategy group led by strong champions and charged with developing protocol, training plans, policy change proposals, etc. Has not yet introduced external partners to TDM plans and sought their involvement in planning. Facilitator positions not created, nor capacity needs identified. Plans for collecting and sharing data from each TDM meeting not yet made; top management provides minimal or no support for design of self-evaluation aspect of TDM. Have not begun discussion with larger agency of icebreaker meetings as a natural result of initial removal and change of placement TDM meetings.	TDM meetings are underway for at least one TDM type (e.g. removals, changes of placement, permanency/ reunification) and/or in at least one geographic area of the site. There is a clear rollout plan for full implementation of all types of TDM across the site. Policy dictates that except in cases of imminent risk, all meetings are held prior to the child's possible move, and always before court. 'Firewalls' have been designed and monitoring mechanisms are in place; intended coverage and proper timing is at least 50%. At least a paper record is kept for each TDM meeting held. Plans underway for collation and sharing of meeting data, with top management support. A plan to work w/ BCP to recruit representatives of the family's own community for TDM participation is in place. Appropriate facilitators (trained, immediately accessible, internal) have been hired and are able to	TDM meetings are held for ALL meetings of one or more of the 3 types. In the alternative, meetings are held for ALL three types in one or more targeted geographic areas. The rollout plan for full implementation is proceeding. Except in cases of imminent risk, all meetings are held prior to the child's possible move, and always before Court. 'Firewalls' are in place and monitoring indicates that intended coverage and proper timing is at least 75%. Representatives of the family's community are invited to meetings of at least one type (e.g. removals) at least half of the time and their participation is tracked and discussed regularly. Facilitator capacity has grown to meet the increased need caused by more meetings. Future capacity needs continue to be assessed and planned for. At least manual tabulations of TDM data are done and publicly shared; top	TDM meetings are held for ALL of at least 2 out of 3 intended situations (removals, COP, perm'y) in ALL the site's geographic areas. Good progress toward the end date for full implementation. Except in cases of imminent risk, all meetings are held prior to the child's possible move, and always before Court. 'Firewalls' working and monitoring indicates intended coverage and proper timing is over 90%. Representatives of the family's community are invited to ALL of at least one type of meeting all the time; there is steady progress toward increased participation which is tracked and discussed regularly. Facilitator capacity continues to grow without any change in high standards for training, accessibility, etc. There is a clear date for full capacity achievement. The site regularly collects and shares TDM data; sophistication of methods	Meetings are being held for all child removals, changes of placement, and permanency decisions, including reunification, across site.. Except in cases of imminent risk, all meetings are held prior to the child's possible move, and always before court. Effective 'firewalls' ensure nearly 100% compliance in intended coverage and timing. Representatives of the family's own community are invited to all meetings, and there is steady progress in achieving full attendance. All meetings are led by trained, immediately accessible, internal facilitators TDM outcome data is collected, widely shared, and regularly used by top management to assess progress and performance and manage change efforts. Every meeting resulting in a child's removal or change of placement also initiates the scheduling of an "icebreaker" meeting.

	fully meet the demand for their services. Assessment of future capacity needs completed and plans for increased staffing are in place. Discussion with RTS about icebreaker implementation has begun.	management and SET utilize this information for self evaluation purposes. As protocol and training are developed for icebreaker meetings, TDM practice supports it by ensuring icebreakers are explained and initiated in removal and COP meetings.	grows; automation explored or attempted; prominence of TDM data in management discussions is apparent, as data are available to describe TDM coverage, results, and the quality of meetings. TDM supports icebreaker practice as in #3.	

Each point on the scale is associated with progress toward implementing several dimensions of work related to the broad strategy. Raters must use their judgment about the site and its circumstances in weighing progress in each area. Generally, choosing a particular point on the scale indicates that the site has accomplished most of the benchmarks described under that point.

Site: _____ Self-Evaluation				
1	**2**	**3**	**4**	**5**
Has not provided an extract of placement data from which longitudinal database can be developed. Has not established a self-evaluation team (SET) representing analysis, data management, and frontline program perspectives. Top management provides minimal or no support for self-evaluation. Little or no experience or staff capability in data management or analysis	Extract has been produced and findings from the resulting database have been shared with SET. SET membership has been identified and meetings are being held. Practice strategy teams are working to produce data describing their progress in putting each strategy in place. Agency management has expressed support for self-evaluation and participates directly or through a deputy in SET meetings.	A plan is in place for regular updating of the longitudinal database. SET has developed a plan for obtaining data about sibling placements. SET has developed a plan for producing outcome data at neighborhood (sub-county) level. Data describing the supply of resource families are being compiled. Data about potential resources families track their progress from application to service. At least manual tabulations of TDM data are done. SET is using newsletter or other media to share info with agency staff and community partners. The director and deputies request and receive regular updates from SET.	Longitudinal database is updated quarterly or semi-annually with breakouts by neighborhood. Data are available about rates of sibling placements. Monthly reports describe the inventory of resource families and success in moving potential resource families into service. Data are available to describe TDM coverage, results, and the quality of meetings. SET is making efforts to link progress in implementing key strategies with changes in outcomes. Management team is using data within the agency and in the community to establish a framework for accountability.	Data are readily available to monitor changes in outcomes and performance relative to key strategies. Data management and analysis capabilities permit ad hoc analysis of emerging needs and trends that have budgetary or policy implications. Agency staff and community partners are aware of key outcomes and agency's performance relative to those outcomes. Self-evaluation is integral to or provides a framework for related performance improvement efforts, such as Continuous Quality Improvement, PIP implementation, or consent decree monitoring activities.

Each point on the scale is associated with progress toward implementing several dimensions of work related to the broad strategy. Raters must use their judgment about the site and its circumstances in weighing progress in each area. Generally, choosing a particular point on the scale indicates that the site has accomplished most of the benchmarks described under that point.

RECRUITMENT, DEVELOPMENT AND SUPPORT OF RESOURCE FAMILIES

1	2	3	4	5
Calls to resource families are not returned, lack of support to RF	Inconsistent support of RF, negative gain in # of RF	Agency recognizes need to support RF and is developing plans	Implementation begins to support RF, positive gain in # of RF	Resource families are well supported and recruit other families
High dependency on shelters, group homes or institutional care, too few RF	Use of emergency shelters/foster homes, children placed away from neighborhood	Closed shelters, children placed in resource families but few in their own neighborhood	Children are placed in their neighborhoods 50% of the time	Children regularly are placed in their own neighborhoods
Non-existent or hostile relationships with neighborhood partners- have not identified potential partners	Neighborhood and agency staff meet to discuss the need for RF and new services.	RFP for neighborhood contract has developed and new services are being developed.	Neighborhood partners have contracts but are unsure how to recruit , develop & support RF	Neighborhood partners are actively recruiting RF & developing services.
RF are not permitted to have relationships with birth parents	RF may have relationships but are not encouraged or supported	RF are encouraged to have relationships with BP, SW uninvolved	Staff, RF and BP have a collaborative relationship in many cases	Staff, RF and BP work collaboratively towards case plan goals
Social workers do not engage resource families; antagonistic relationships w/comm.	Few Sup. and SW understand importance of relationship with RF & Community.	Sup. trained to recognize the importance of RF & Community.	Sup. holds staff accountable for their relationship with RF & the Community	All staff, private partners and the community support RF
The agency does not consider resource families a valuable resource	The agency recognizes the value of RF but is paralyzed to act	The agency begins to review/revise policies and personnel practices to reflect the valuing of RF	All staff appreciate the valuable services that RF provide to children and their families	All staff are evaluated on their support & promotions based on level of support/appreciation of RF
Resource families are excluded from policy development and placement decisions	Some staff & RF recognize and request more input in policy development & placement decisions.	Staff recognizes the importance that RF can provide in placement decisions and policy development and develops plan to allow this to occur.	Agency assists RF in being able to participate in placement decisions and policy development.	RF are expected to participate in policy development and placement decisions
Training is inaccessible, uninspiring and outdated. No use of teens or parents in class.	Training is held at various times but at agency. More lecture & review of rules than involving activities; rigid training rules	Training held at various times/locations, adheres to adult learning theory, uses many techniques/strategies	Training begins to include youth, birth parents and resource families	Training is held in the community, is exciting, vibrant and relevant to parenting the children in out of home care

Site: _____ Building Community Partnerships				
1	**2**	**3**	**4**	**5**
Agency Director and key management staff are being orientated to values, core strategies, and key elements of F2F and the idea of partnering with communities for child safety and permanency.	Agency Director understands and values the partnership with neighborhoods and develops a strong internal message to support this belief. Director's message is systematically incorporated into management and supervisory meetings related to the value and expectations of community partnerships.	All agency staff understands and embrace F2F strategies and values of community partnership. Child welfare activities begin to move out into the communities of focus. Geographic assignment of cases is explored.	Visitations, TDMs, Agency Departmental and/or supervisory meetings take place in priority communities regularly. Direct line staff is familiar with and known in the community.	Agency staff is reviewed through performance evaluations that include F2F values and activities. Staff is promoted accordingly.
BCP workgroup has not been established and champions have not been identified	BCP work group is established and chaired by strong champions that understand the value of neighborhood based partnerships. The work group has identified potential community partners and invited them to participate in regularly held meetings.	Goals of the BCP work group are outlined in regularly reviewed /measured through strategic or action plan formats.	Neighborhood contracting and collaboratives are established. Neighborhood based work impacts the outcomes of children and family through a continuum of care.	Neighborhood contracts and collaboratives impact child welfare outcomes. Collaboratives are viewed as infrastructure of neighborhood based work. BCP workgroup serves as a monitor to community based work while geographically assigned managers and staff participate in collaboratives to set new goals.
Data on neighborhoods with the highest # of kids in care has not been identified .	Preliminary data has been identified re: neighborhoods of priority; numbers of children in care from neighborhoods; number of foster homes in priority neighborhoods; number of private provider homes in priority neighborhoods; number of children placed in their own neighborhood.	Regular community forums, events, town hall meetings are held to share the neighborhood based data and change in practice. Agency Director is delivering the message and is available to the community for this dialogue.	Community sees itself as a partner with the agency and is working collaboratively to deliver data and message of agency. Data is provided to community on regular basis regarding neighborhood children.	Neighborhood based child welfare activities are measured and impact child welfare outcomes. Data is understood and owned in partnership with child welfare agency.
A job description for the F2F Coordinator	Agency Director creates a position for F2F Coordinator	Community partners and BCP workgroup are coordinating	Community partners and private providers are	Community partners participate in all four core F2F strategy work

has not been developed	and selects an individual whom understands and values community partnerships and has a strong connection to neighborhood based work. Agency Director and management staff work closely with F2F Coordinator to oversee all strategy groups.	existing neighborhood resources through the development of neighborhood resource guides and are identifying gaps in services to support PCWA workers as they work with families. Private Providers are also engaged to work closer with the community and the PCWA	developing strategies and services to address the needs of neighborhood children based upon data through collaborative partnerships.	groups. F2F Coordinator and management/steering committee have made systematic connections and changes in policy to support the ongoing progress of the building of community partnerships.
No community representation in TDMs is occurring.	TDM strategy is introduced to community and initial recruitment of community reps begins in partnership with TDM workgroup.	Community reps have been oriented and begin to attend TDMs for neighborhood families. An internal mechanism has been developed to notify the reps when TDMs have been scheduled.	Community reps are connected to the neighborhood collaboratives to coordinate and link families to neighborhood supports through collaboratives continuum of care. Community reps are attending removal TDMs and prioritizing other critical TDMs.	Community reps attend 100% of removal, disruption, and reunification TDMs. Families are linked to ongoing supports through collaboratives.
No neighborhood based foster care is available or identified.	An initial recruitment plan for neighborhood based foster care is developed in partnership with RTS workgroup.	Neighborhood based foster care recruitment and training is occurring in priority neighborhoods. Internal placement process is being assessed to initiate kinship care and neighborhood placement as a priority	Neighborhood based foster homes are available through agency and private providers where neighborhood children are placed in their community of origin. Neighborhood based resource families have support groups in neighborhood and are supported/connected to their neighborhood's collaboratives.	Together neighborhoods, PCWA, and Private Providers ensure Child safety and permanence. This is an on going effort. Creating ever widening circles of Influence, opportunity, and funding Possibilities.

Each point on the scale is associated with progress toward implementing several dimensions of work related to the broad strategy. Raters must use their judgment about the site and its circumstances in weighing progress in each area. Generally, choosing a particular point on the scale indicates that the site has accomplished most of the benchmarks described under that point.

Appendix E

Key Elements of the Core Strategies

KEY ELEMENTS OF THE BUILDING COMMUNITY PARTNERSHIPS STRATEGY

GOALS

- To build authentic partnerships between the public child welfare agency (PCWA) and community residents, service providers, licensing agencies, resource families to work together with families and youth for improved family outcomes.
- To enhance the capacity of these partners to address issues impacting families and their ability to care for their children.
- To utilize the strength of the community-PCWA partnership to eliminate racial disproportionality and disparity in all areas of child welfare.
- To identify and utilize neighborhood services and family strengthening supports that are accessible, affordable, culturally connected and geographically available.
- To recruit and establish a strong network of neighborhood-based resource families that are trained and supported to work with children, youth, and their birth parents toward reunification and other permanency goals.
- To ensure that parents and youth have community representatives at TDM meetings to serve as an advocate for families and youth.

- VALUES

- Children and youth belong in their own families.
- All communities have strength that should be accessed by the PCWA in support of children and families.
- The work of the PCWA and its contract agencies are strengthened when they reach out to communities and invite their authentic participation.

- ASSUMPTIONS

- The community is vested in the well being of all of it's members and to that end will contribute to their health stability and well being of all it's members

- Vulnerable Families and youth involved with child welfare are entitled to support and services and benefit most when their communities, the public systems, service providers, and foster care agencies work in partnership.

- Children and youth whose parents are unable to keep them safe have a right to have extended family be their first placement option and to stay connected with their parents, siblings extended family and community. by the child welfare system.

- When it's not possible to place children and youth with their kin, they should be placed in the least restrictive environment consistent with their needs, background, shared values, school placement, and places of worship and recreation. Neighborhood based family foster care provides this critical resource when needed.

- Kinship, foster and adoptive families are entitled to support and resources from their communities and neighborhoods and from the public child welfare agency. These supports to resource families increase their skills and abilities to care for the children and youth in their care. Working cooperatively with the family increases the chances for reunification and permanency between the children and their parents.

- Children and youth placed in more restricted environments have the right to stay connected to their parents, siblings, kin and extended families to resolve issues and reunify them with their families. Custodial care of children must be targeted, focused and time limited. It is a critical intensive intervention and not to be considered permanent. If parent's rights have been severed then targeted recruitment to find them a forever family is incumbent on the community partnership who must work uniformly and expeditiously on their behalf.

- The racial disproportionality and disparity evident in child welfare and other systems can be reduced and ultimately eliminated by a strong and effective collaboration between the PCWA and communities that have suffered the most adverse impact.

- KEY ELEMENTS

- The PCWA acknowledges values and commits to an authentic partnership with the community.

- The PCWA commitment requires strong focused leadership; a trained and skilled workforce, a sustainable infrastructure, and shared accountability from all parties.

- All levels of the PCWA--leaders, managers, supervisors and workers along with foster care agencies and community service providers--must develop a unified vision to assess their current organization and practice to set targets for improvements. This could include zone-based or geographic assignment of staff.

- The PCWA dedicates staff such as a Family to Family Coordinator, community liaison or other champions and resources to this vision and the meetings necessary to build it.

- The PCWA along with the community will identify key leaders, stakeholders, informal and formal neighborhood organizations to forge a cohesive collaborative of services and resources, to provide an array of ongoing supports to birth parents, resource care providers and youth in care.

- The PCWA will form a work group which includes community partners as part of their infrastructure. It is made up of representatives of the community and from the other F2F strategies who meet consistently and frequently for visioning, creating, implementing and evaluating their unified efforts toward better outcomes.

- By the PCWA sharing and discussing baseline data about the nine F2F outcomes and developing tracking tools to capture interventions, activities and results the agency and

community will learn about the child welfare experiences of neighborhood children and families. This helps the partnership focus on realistic benchmarks for change.

- The PCWA and placement provider agencies' recruitment staff will work with the community partners to help identify and support neighborhood residents, who provide kinship care and serve as resource families.

- The PCWA provides funding opportunities to neighborhood based community partners for the implementation of activities such as prevention and reunification; wrap around services for parents and youth; TDM commitments and participation; visitation services; recruitment and support of resource families, as well as other innovative practices that they agree to do.

KEY ELEMENTS OF THE RECRUITMENT, DEVELOPMENT, AND SUPPORT STRATEGY

GOAL

To assure that children who cannot be protected in their own home will be placed with a safe and stable family member or with a family from their own community or neighborhood.

VALUES

- Every child deserves a family for protection, nurturance and permanency.
- Every family needs the support of their community.
- Public Child Welfare Agencies (PCWA) need community partners.

ASSUMPTIONS

1) Racial/ethnic disparity and disproportionality exists in child welfare systems. In order to positively change outcomes for all children, disparity and disproportionality must be addressed at the policy, programmatic and practice levels.

2) Resource families provide the most valuable service we have to offer children in out of home care. These families need the support of all Public Child Welfare Agencies (PCWA) staff, community partners and contract service providers in order to provide protection, stability and nurturance to these children.

3) There are strong, nurturing family and community members who will step forward to care for those children who cannot be safe in their own homes if we inform them of the need and provide them with support.

4) Birth families, youth and resource families are experts on their own experiences and can serve as valuable partners in recruitment, training and developing resource families, PCWA staff and contract providers.

5) Data drives the RDS process, providing information about the children in out of home care and information about how the children and resource families move through the system so that practice adjustments can be continually made.

6) The most successful recruitment efforts target potential caregivers for those children most often left behind (e.g. children of color, teens and sibling groups) as well as the neighborhoods with the highest rate of Child Welfare reports. This mode of recruitment is personal and leaves a residue of goodwill with the community. General awareness media campaigns may attract more attention but there is little evidence that they result in significant gains in the number of foster and adoptive families. Media campaigns can best be used to inform the community of the needs of children in care and identify the areas of greatest need.

KEY ELEMENTS

1) Strong, knowledgeable PCWA leaders model and promote a culture of respect, support and empathy for current and prospective resource families. Agency policies, staff evaluations and the criteria for promotions reflect this culture.
2) All PCWA staff understand that *Recruitment is Everyone's Business and Support is EVERYONE'S Job!*. They know and show that recruitment begins with respect and support for the current, prospective and re-recruited resource families.
3) The Recruitment, Development and Support activities are community based and located in areas where children are most often removed to ensure that recruited families reflect the race, ethnicity and culture of children coming into care.
4) The PCWA staff works with neighborhood partners, youth, Birth Families and Resource Families to establish a well developed system of neighborhood based family care to meet the needs of all children coming into care, including teens and sibling groups.
5) Placement services and supports offered to families are based on the needs of the children and strive to eliminate the racial/ethnic disparities associated with disproportionality. The placement process is conducted in a manner that is respectful and supportive of the children and families involved.
6) Resource families, social workers, youth, community partners, and birth families work as a team in every step of the process leading to reunification or another permanency goal. They participate in placement decisions at TDM meetings; they have an opportunity to play a major role in recruitment, training, and placement and they serve as active members of policy development work groups.
7) A meeting between the caseworker, resource family, and birth parent is held within three days following all new placements or any change of placement to introduce the two sets of parents and to give them an opportunity to exchange any information that will benefit the child. Youth can also participate in and contribute to an icebreaker. These meetings have different names in different locations (e.g. Icebreakers or Family Team Meetings) but the goal is always to begin to build a bridge between the birth and resource families for the good of the child.
8) In order to assist the resource families in caring for the children placed with them, all child-placing agencies practice full disclosure by securing and providing to those resource families all information about the child in their care prior to and throughout the time the child is with them.
9) The PCWA takes a ongoing leadership role in partnering with contract agencies involved with all aspects of RDS work in order to implement the values, goals, and strategies of Family to Family especially in the elimination of disparity and disproportionality across all outcomes.
10) Training for new and experienced resource families is relevant, participatory and accurately describes the experience of resource families. It is accessible geographically and culturally. Foster youth, birth and resource families function as co-trainers sharing their experiences and their lessons learned.

KEY ELEMENTS OF THE SELF-EVALUATION STRATEGY

GOALS

- To provide a comprehensive baseline perspective on agency performance using the outcomes emphasized in Family to Family.

- To create a capability for determining how changes in agency policies and practices, especially those related to key Family to Family strategies, are affecting those outcomes.

VALUES

Self-evaluation is:

— focused on outcomes;
— carried out by a diverse team of child welfare managers and staff, neighborhood partners, and private providers; and
— supported by the investment of human and other resources in acquiring technical expertise in analysis and data management.

ASSUMPTIONS

- Longitudinal data that capture the experiences of all children served by the child welfare agency provide the most reliable estimates of where the agency stands on the outcomes emphasized by Family to Family.

- The responsibility to use data to understand agency and provider performance is shared by all staff and is not just the job of data managers and analysts.

KEY ELEMENTS

Self-evaluation seeks to enhance data resources and to develop an ongoing process for using data to help make policy and practice decisions.

Data Resource Key Elements:

- Develop within 6 months a longitudinal database that describes baseline performance relative to Family to Family outcomes;
- Establish a process for updating the baseline database every 6 months after the initial baseline is established;
- Develop indicators that allow analysis of whether: (1) siblings are placed together; and (2) children are placed in their own neighborhoods; and
- Work with strategy teams to develop process data that track the implementation of key strategies, with appropriate benchmarks identified for each strategy.

Self-Evaluation Process Key Elements:

- Establish a self-evaluation team (SET) representing three key perspectives—frontline (agency staff and community partners), data management, and analysis;
- Develop a concrete and well-defined linkage between the SET and top management so that managers act to reinforce self-evaluation; and
- Establish a regular schedule for SET meetings and adjust agency participants' assignments to accommodate their commitment to self-evaluation efforts.

KEY ELEMENTS OF THE TEAM DECISIONMAKING STRATEGY

Goal

To involve birth families and community members, along with resource families, service providers and agency staff, in all placement decisions, to ensure a network of support for the child and the adults who care for them.

Values

- Every child deserves a family
- Every family needs the support of the community
- Public child welfare agencies need community partners

Assumptions

- A group can be more effective in decision making than an individual.
- Families are the experts on themselves.
- When families are respectfully included in the decision making process, they are capable of identifying and participating in addressing their needs.
- Members of the family's own community add value to the process by serving as natural allies to the family and experts on the community's resources.

Key Elements

1. A TDM meeting, including birth parents and youth, is held for ALL decisions involving child removal, change of placement, and reunification / other permanency plan.
2. The TDM meeting is held *before* the child's move occurs, or in cases of imminent risk, by the next working day, and always before the initial court hearing in cases of removal.
3. Neighborhood-based community representatives are invited by the public agency to participate in all TDM meetings, especially those regarding possible child removal.
4. The meeting is led by a skilled, immediately accessible, internal facilitator, who is not a case-carrying social worker or line supervisor.
5. Information about each meeting, including participants, location, and recommendations, is collected and ultimately linked to data on child & family outcomes, in order to ensure continuing self evaluation of the TDM process and its effectiveness.
6. Each TDM meeting resulting in a child's removal serves as a springboard for the planning of an "icebreaker" family team meeting, ideally to be held in conjunction with the first family visit, so that the birth-foster parent relationship can be initiated.