Economic Incentives, Adoption out of Foster Care, and Long Term Outcomes*

David Simon
Aaron Sojourner
Jon Pedersen
Heidi Ombisa Skallet

July 11, 2022

Abstract

We investigate how financial incentives to adopt foster children impacts a range of child outcomes. We leverage rich administrative data from Minnesota and a unique policy change that raised payments in permanency to eliminate the disparity with payments in foster care. Equalizing the subsidy between adopted and foster children, decreases the length of foster care episodes. There is no evidence that the "marginal" adoption caused by the more generous stipend is less stable. Continuing to provide the full foster subsidy in the years post adoption increases a child's school attendance, their likelihood of staying in the same K12 school, and substantially raises their standardized academic achievement scores. We further show interesting dynamics in how a more generous subsidy post adoption impacts children's utilization of mental health services. This paper adds to a small body of work investigating how financial incentives for adoption decreases the time spent in foster care. It is the first paper to use casual inference methods to connect such financial incentives to long term child outcomes.

*Simon: University of Connecticut & NBER (david.simon@uconn.edu). Sojourner: University of Minnesota & IZA (asojourn@umn.edu). Pedersen & Ombisa Skallet: Minnesota Department of Human Services. Sojourner acknowledges funding from Casey Family Programs.

1 Introduction

Financial incentives for adoption out of fostercare have potentially large implications for child wellbeing. Up to 6% of children will experience a foster care episode by the time they turn 18 years old (Wildeman and Emanuel, 2014). While most of those children will reunite with their parents, many will remain in foster care for a prolonged period of time, and 10 to 25% will age out without being re-unified or adopted. Prolonged exposure to foster care is associated with adverse outcomes, and child welfare agencies struggle to find safe, supportive care for foster children. Foster children are the state's responsibility and caring for children requires many expenses so agencies pay foster caregivers a monthly stipend. Foster parents face a dynamic choice each period to either: continue to foster the child, take permanent legal custody over the child (such as through adoption), or to discontinue fostering. Such a decision will be a function of the utility from fostering, having an adopted child, and the financial costs of providing either foster care or a permanent home.

When a child transitions out of foster care and into a permanent living arrangement, states tend to diminish or completely stop payments, requiring the permanent parents to take on full financial responsibility for the child. A reduction in financial support in the move from foster to permanent care can inadvertently create an incentive that discourages permanency and extends children's time in foster care. It could also affect match quality between child and parent in complex ways. The highly personal nature of child rearing means personality matching between prospective adoptive parents and children is important and it remains unclear how other incentives affect parent/child match. On the margin, such incentives could overpower intrinsic incentives and worsen match quality, with negative consequences for the child (Bowles, 2016). Additional financial resources coming into the household might themselves improve child outcomes. Further, it might enable low-income family members who have the child's interests at heart to take on the responsibility of care rather than having to rely on potential adoptive parents with more private resources but less connection to the child's community of origin.

We leverage a unique policy change in Minnesota that lifted payments for adoption support to equal foster care payments for foster children age 6 and above. The policy change is unusual in equalizing the stipends offered in foster care and in post-foster, permanent arrangements, adoption and transfer of permanent legal and physical custody to kin (TPLPC), potentially resulting in much bigger changes in resources and larger impacts than earlier work. Thanks to linked administrative data and special data work enabling post-adoption linkages despite any child name changes, we follow children over time between foster, adoptive, TPLPC, and original family settings, we can look at some outcomes not previously investigated in the literature, such as child academic achievement and the likelihood of re-entry into foster care, a measure of the quality of the post-foster placement.

The policy experiment we study, Minnesota's Northstar reform, was inspired by concern about the disincentive created by lower payments in permanency. Its implementation was preceded by a demonstration study with experimental and non-experimental designs that found some promising results.[1] Northstar was designed to scale this demonstration project up statewide and accelerate foster care exit by reducing the disparity in stipend between foster and permanent care. Motivated by the fact that younger children exited into permanency relatively quickly already, it treated older and younger children differently. For foster children age 6 or older when entering permanency, the reform eliminated the payment disparity, equalizing stipends in foster care, adoption, and permanent kin care. For foster children below age 6 when entering permanency, the disparity between foster and permanent care was not eliminated. Payments in permanency were set at half those in foster care for these younger kids. So the reform had a much larger effect on caregiver financial incentives for older children than for younger. The Northstar reform had other aspects, such as shifting to a new instrument for assessing child and family needs and resources to determine payment levels, but

[1]The demonstration project aimed, "to determine whether a continuous (or single) benefit program would increase permanency rates and shorten foster care stays among children who have been in foster care for an extended period of time. The continuous benefit program raised the public assistance benefits received by caregivers who adopt or accept permanent legal and physical custody of their foster children to a level equal to the rate paid for foster care. Historically, benefit rates paid to eligible caregivers who offer a foster child a permanent home could be half the rate paid for foster care." (of Applied Research, 2011). The study suffered from a variety of implementation challenges. An experimental part of the study ran in 2 counties and found acceleration into permanency resulting in 10% (85 days) shorter foster episodes, without changes in the distribution of permanency types (adoption, permanent guardianship, family re-unification...).

none built in a structural differentiation at this age threshold or other age thresholds. We study whether the demonstration project's estimated benefits on reduced time in foster care replicated in the scaled-up policy and, using linkages across state agencies' administrative data, extend the analysis to outcomes beyond the child welfare system into longer-run outcomes in schools and the community: test scores, school attendance, school stability, and mental health cases.

To understand if Northstar's changes in financial incentives affected children's length of time spent in a foster care spell, their probability of exiting into adoption or permanent kin care, and the stability of permanent placement, we use a difference-in-differences logic. The change in outcomes experienced by younger foster children after versus before the Northstar reform captures all of the following: the effect of a policy change which diminished but did not eliminate the disparity, other associated changes in the child welfare system such as assessments instruments, and any secular changes in the environment that would affect outcomes. However, the change in outcomes experienced by older foster children includes all of that plus one additional factor – payments generous enough to completely eliminate the stipend disparity. We contrast outcome changes experienced by older foster children against changes experienced by younger foster children to isolate effects of payment incentives. Although all the analyses share this difference-in-differences design at the core of their identification approach, we use different auxiliary assumptions to model different kinds of outcomes. To understand impacts on the probability of different types of exits from foster care (adoption, kinship care, reunification...), we use duration and linear probability models. To understand the impact of the policy on longer-run outcomes such as standardized academic achievement scores, attendance rates, and propensity to switch schools in K12 and probability of receiving state funded mental health services in the community, we have to deal with the fact that the time the child spends in foster care and, hence, their exposure to alternative policy regimes may be endogenous. To deal with this, we predict foster spell length for each case based on characteristics fixed at case start and use predicted exposure to policy based on this as the treatment variable.

While the majority of children taken into foster care will be reunited with their families, many will be separated permanently and the children will remain in foster care until they are either adopted either by foster parents, relatives, or they age out of the foster care system. While prior descriptive studies show a relation between pecuniary incentives and adoption, many are plagued with selection bias and there study how subsidies impact the move to permanency. 10 to 25% of children will age out of the foster care system without being reunited with their families or adopted. Aging out of foster care itself is correlated with poor transitions into adulthood including homelessness: with up to 43% of children who transition out of foster care spend time homeless by the age of 21 (Service (2019)). It is fiscally expensive for states to provide the financial and administrative support to their wards, such that there could be net monetary gains to continuing to offer stipends post adoption.

Historically, the quantitative analysis of foster care policies is likely subject to selection bias (Cuddeback, 2004; Buckles, 2013), More recent studies trying to get at causal impacts have focused on the question of whether home removal is good for child well being.[2]

A related literature evaluates how monetary incentives for foster care impact foster families and time spent in foster care. Testa and Slack (2002) and Doyle (2007) investigate this in the context of a policy in Illinois that decreased payments to foster care parents. These papers find that as payments to foster care families decline both the amount and quality of care declines: particularly for "costly" child cases such as children with diagnosed mental health problems. Likewise, Testa and Rolock (1999) and Doyle and Peters (2007) show that higher payments to foster care families increase supply of available foster families. This part of the literature also shows improvements in placement stability and quality of foster care increases with increased payment levels. Finally, Buckles (2013) look at federal funds provided for adoption subsidies to show that such funds can increase the number of adop-

[2]Two influential papers use random assignment to case workers and variation in relative case worker leniency (Doyle Jr, 2007, 2008). Doyle compares children who were assigned to caseworkers with a predisposition for home removal to children assigned to a more lenient case worker. Applying this design to Illinois data: Doyle finds worse outcomes for children (particularly older children) who were removed from the home. More recently, Roberts (2019) and Bald et al. (2019) use data from other states to implement a similar research design and find improvements in educational outcomes for younger children, but null effects of home removal for other children and outcomes. Warburton et al. (2014) document a sudden increase in home removals following a high profile child death due to a failure in the foster care system. Such removals are associated with worse educational and economic self sufficiency outcomes for older boys.

tions; while Argys and Duncan (2013) uses a state-year panel to look at the impact in variation in state subsidies on number of adoptions.

We contribute to the earlier literature by showing that increasing the state adoption stipend to eliminate the disparity in payments results in a large increase in the hazard rate of foster care exits for older (relative to younger) children. This is the first paper we know of to specifically focus on the complete elimination of the disparity in the stipend, with larger impacts than found in earlier papers on financial incentives for adoption. Second, we are the first paper we know of to look at outcomes related to child human capital attainment: school attendance, likelihood of staying in the same school over a year, child academic achievement, and utilization of mental health services. Earlier work could only speculate that financial incentives to adopt were ultimately beneficial to the child. Skeptics could argue that the "marginal" adoption caused by financial incentives could be less beneficial or even harmful to a child. However, we find improvements in academic outcomes that are enough to eliminate the negative correlation in achievement associated with being placed in foster care.

2 Context and Data

After removing a child from their family of origin, state child welfare agencies take responsibility for the child and place the children in care with foster parents or an organization that provides these services. Generally, agencies intend to find any foster child a permanent, private family with foster care as a temporary, transitional arrangement. Leading options for permanent arrangements that might end the need for foster care include (1) reunification with the family of origin, (2) kinship guardianship arrangement to someone with pre-foster-care relationship with the child, such as a grandmother, uncle, or family friend (called transfer of permanent legal and physical custody, or TPLPC, in Minnesota) (3) adoption, whether by the child's relatives, kin, current or former foster parents, or a new adoptive resource family, (4) aging into adulthood and out of foster care, and other less-common possibilities.

Starting in January 2015, the state of Minnesota adopted a set of child welfare reforms that included aspects of payment equalization between foster and permanent care arrangements to eliminate the financial disincentive, shorten children's time in foster care, and increase the quality of permanent care. For foster children age 6 or older, the reform equalized payments whether the child was in foster care, permanently adopted, or in a permanent TPLPC kinship arrangement. This new commitment to permanent payments until children reached adulthood required a significant commitment of new public resources because previously any payments that continued into permanency (potentially due to a child being eligible based on having a disability) were considerably lower. Foster and permanency payments were reformed but not equalized for foster children below age 6, who already had higher likelihoods of exiting foster care into permanency than older foster children. For young foster children, payments in permanency were set at half the rate of payments in foster care. In most cases, the reform increased payment levels in permanency for young children but a large gap remained between payments in foster and permanent care.

Additional reforms occurred around the same time as these payment reforms, which demand careful attention so as not to confuse their effects for the payment reform effects. First, the state adopted a new system for assessing children and caregivers and determining payment levels. The change had two elements. One, the assessment instruments themselves changed so that those used in the earlier period were not used in the later period. Two, it moved away from a system in which all children were assessed in a rubric that determined foster payments but only subsets of children were assessed using different rubrics to determine potential payment levels in permanent arrangements to a new system in which all children were assessed in a single rubric that determined payments in both foster and permanent care.

Second, a few months after the Northstar payment and assessment reforms, the state added a requirement that TPLPC (Minnesota's version of kinship and guardianship care) could occur only after a licensed foster parent had cared for the child for 6 months. This was part of the state's decision to opt into a federal guardianship assistance program and the additional licensure requirement was required in order to improve placement stability following permanency, and ensure the child and their permanent caregiver had a strong attachment to one another. Finally, children were less likely to qualify for eligibility for payments in permanency prior to the Northstar reform: specifically, there were greater requirements related to the child being disabled or being adopted into a home with a disabled sibling that was relaxed post reform.

Unlike the payment reforms, neither of these other coincident reforms introduced a sharp discontinuity in policy between ages 5 and 6. Therefore, our analysis will focus on comparing changes among foster cases before versus after the reform with special attention to any differences in changes for children above and below this age threshold. All kids were subject to coincident reforms. Younger kids had incomplete payment equalization while older kids had full equalization. A difference-in-differences design interprets the difference in changes between the older and younger groups as informative about the impact of full versus incomplete payment equalization.

2.1 Foster Care Data and Descriptive Statistics

The state of Minnesota provided records on the 54,577 foster care cases where a child entered foster care starting between January 2011 and July 2019 [3]. Each is called a continuous placement episode (CPE). Panel A of Table 1 shows summary statistics for the characteristics of each CPE. Just over half started after the Northstar reform. During a CPE, the foster child may move across a sequence of placements, such as a temporary group home and then a foster family. For each CPE, we observe its start date, a unique

[3] there were also a small number of cases where the parents voluntarily placed their children in fostercare. Because these are nearly guaranteed to end in reunification we excluded them from our analysis

but anonymous child identifier that allows linking of the same child across multiple CPE, child demographics (age, race/ethnicity, gender), reasons for removal from the child's family of origin, the sequence of moves between particular foster care placements within the CPE with each placement's start month and end month (if placement ended by July 2019), and the CPE's end month and permanency disposition type if the CPE ended by July 2019. Children average just under age 8 at the start of their CPE and range from 0 to 18 (Table 1). The observed CPE length averages 11.4 months (calculated conditional the full CPE being observed). On average, children with any CPE starting in this time period averaged 1.37 CPE starting in this time period. Thirty percent of CPE involved African-American children and 25 percent involved Native American children, far higher than their share in Minnesota's child population.

We partnered with the Minnesota Linking Information for Kids (Minn-LInK) data lab, which facilitates confidential linking of state administrative data sets. Matching is done using several software packages to implement probabilistic matching which is then verified by hand to ensure accuracy. This enabled a longitudinal view of each foster child's path through the child welfare system linked to their experiences in any Minnesota public school, including data on student attendance, standardized achievement tests starting in grade 3, and school disciplinary records.

This is the first study to benefit from a recent effort by the Department of Human Services to link records for the same child over time in cases when adoption led the child's name to change. This enables better measurement of post-adoption outcomes such as student test scores and re-entry to the child welfare system. The latter is an important measure of the quality of the post-exit arrangement.

Any child's prior experience with the child welfare system might be informative about the likelihood they remain longer in foster care. These differences between children are important predictors of outcomes. Using each child's past case history, we control for the number of times a child was in fostercare. On average, a 10 year old entering foster care would have been in foster care once before.

A primary outcome of interest is probability of exit from foster care and into various permanent arrangements. In theory, raising payments in permanency will reduce disincentives to delay permanency and raise the monthly probability of exit from foster care. Because the payment reforms boosted payments in permanency for TPLPC and adoptive caregivers, not for reunification with family of origin or children aging out of foster care or other dispositions, we would expect an increase in the share of case going to TPLPC or adoption and decreases in other dispositions as well as just an acceleration of exit into permanency with reduced time in foster care. Because these payment reforms most strongly affected children age 6 and older rather than those 5 and under, we would expect stronger effects in the older group. We observe the disposition of each foster spell in 88 percent of cases and 12 percent of spells continue past our observation window's end in July 2019.

Reasons that the spell ends are either reunification with family of origin (58 percent), adoption (12 percent), TPLPC with kin (8 percent), aging out of foster care (2 percent), and all other reasons (10 percent). For children with cases starting in this time range, we can usually see prior cases as well.[4]

To investigate the impact of the policy on time to exit from foster care: we use hazard models where the outcome is whether an exit is observed in a given month. In this case, relevant predictors (age and calendar year) change at the monthly level and the outcome variable (whether or not there is an observed exit) changes at the monthly level. We transform the data so each observation represents a CPE-month. We allow age and calendar time to vary within each CPE spell so that we could pinpoint when children would progress into different policy regimes. Panel B of Table 1 shows summary statistics for this CPE-month level data. When structured this way the data has 699,413 child-month observations with 54,577 different CPEs from 43,633 children.

3 Identification and Estimation

To study how the shift to the Northstar policy regime impacted time to permanency and child outcomes we use a variety of difference-in-difference type specifications. Exogenous variation comes from comparing changes in average outcomes before the reform to after for children below 6 (comparison)

[4]All past records on a child are preserved for ten years after the child's most-recent case ends.

versus those 6 and older (treated). For child-i observed in time t at age $a(it)$, consider the following difference-in-differences model:

$$Y_{it} = \beta_1 1(post)_t 1(age6+)_{it} + \beta_2 X_{it} + \gamma_{a(it)} + \delta_t + \epsilon_{it} \tag{1}$$

Y_{it} reflects i's outcome at month t. X_{it} is a vector of control variables such as child demographics and indicators for the reasons for the child's removal from foster care [5]. γ_a and δ are years-of-age and calendar year-month fixed effects. Our interest centers on the parameter β_1, which expresses how much more average outcomes change for kids in the older age group after reform than average outcomes change for kids in the younger age group. This is a non-staggered difference-in-difference design estimating a single policy change where the age and time fixed effects subsume indicators for "treated" and "post" respectively.

Because we focus on contrasting changes across the sixth birthday threshold, we consider 3 different analytic samples: children of all ages, those 2-9 only, 4-8, and those 3-7 only. There is a bias-variance trade-off: focusing on the narrowest range of ages and excluding cases far from the threshold keeps those on either side most similar, which reduces bias, but also reduces sample size, which increases variance. Including cases farther from the threshold potentially increases bias while also reducing variance. In addition to a bias-

[5]We show results with and without controls in our main model. The complete set of controls include: race (white, African-American/Black, Native American, Asian, Pacific Islander, Unknown, and other), Hispanic Ethnicity, reason for removal (neglect, physical abuse, care taker drug use, behavioral problems, and other), gender, and a child's total number of foster care placements.

variance trade off this exercise also changes the local average treatment effect because the increased subsidies might impact the older children more than younger children.

The identifying assumption in this model is that on average unobserved influences do not change systematically with the policy,

$$E[d\epsilon|d1(post), 1(age6+), X, 1_t] \equiv E[d\epsilon|1(age6+), X, 1_t]$$

Aside from the payment policies described above, no other policy changes had an age-break at or near age 6.

We perform event studies to make explicit any differential trends between treatment and control groups and to ensure that any changes in outcomes ascribed to the policy occur at the time of the policy. We additionally show placebo tests on younger children and the impact of the policy on observed covariates as additional tests of this assumption. [6].

3.1 Estimation: time to permanency

The intended purpose of the Northstar policy was to decrease the time a child spends in foster care. To evaluate its success; we adapt equation 3.3 above so that each observation is a month that a child was observed in a continuous

[6] There have been a number of recent critiques of two-way fixed effects models (Goodman-Bacon, 2018; Baker et al., 2021). However, these focus on staggered-adoption designs, where there are with multiple treatment and control groups whose timing occurs at different times and with unbalanced pre-and post periods. We do not have that issue given there is a single treated group and all treatment began at the same time

placement episode. Y_{iat} is an indicator variable equal one if the child exited foster care that month and zero otherwise. In some specification this will reflect any exit to permanency and other times a specific type of exit such as adoption or reunification. One problem with estimating equation 3.3 above is that some of our outcomes, such as the probability of exit in a given month, is a function of the duration of the foster care spell. Without accounting for this duration dependence, estimates of β_2 could potentially be biased due to a mechanical correlation between time spent in foster care and increased likelihood of exit. We employ different survival models to deal with this issue. The most common type of survival model is a Cox proportional hazard, this has previously been used in difference-in-differences models on foster care spells (Buckles (2013)). Consider survival time modeled as follows:

$$h_{iat,p|\mathbf{x(p)},\beta} = h_0(p)e^{\mathbf{x'}\beta} \tag{2}$$

Where $h_{iat,p|\mathbf{x(p)}}$ is the proportional hazard of exit in period t. Likewise, p in the above equation signifies duration of the continuous placement episode, for which the hazard accounts. We estimate the difference in differences by substituting the right hand side of the linear model in 3.3 into $\mathbf{x(p)'}\beta$. Taking logs of both sides of 2:

$$\ln(h_{iat,p}) = a(p) + \beta_1 1(age6+)_i + \beta_2 1(post)_t 1(age6+)_i + \beta_3 X_{it} + \gamma_t + \epsilon_{it} \tag{3}$$

where now, h_{iat} is a hazard of different exit types (any, TPLPC, adoption, or reunification), β_2 is the difference-in-differences hazard ratio: the % differences in likelihood of exit to h_{iat} for treatment relative to control groups. $a(p)$ is now the baseline hazard: the likelihood of exit as a function of duration of the CPE; which Cox estimates without the need for an assumption about functional form. Standard errors are clustered on child. A benefit of Cox proportional hazards is that they are easy to interpret, implement, and are commonly used in survival analyses. However, the Cox model assumes that the hazard functions across groups are proportional over the duration of the spell. Such an assumption could be violated in a difference-in-differences analysis if within a continuous placement episode there is a change in exit probability of the treatment relative to control group over the duration of the spell (i.e., the two groups are not proportional in exit probabilities conditional on having been in the spell for the same variation). However, it is worth noting that variation in our research design happens both within and between continuous placement episodes. To the degree that probabilities of exit between treatment and control groups shift across the entire CPE, this will not bias the results. The Cox proportional hazard model also has the benefit of not requiring the econometrician to make any assumptions about baseline hazards.

As an alternate specification we use a discrete-time proportional hazard model with a complementary log-log form. This trades off having to make assumptions about the functional form of the baseline hazard in exchange for fitting a discrete time model. The log-log form is the simplest functional form. Specifically, we make the following assumption about the form of the hazard:

$$h_{iat,p|x(p),\beta} = 1 - e^{-e^{r_{iat}(p)'\beta}} \tag{4}$$

We substitute a modified version of the linear model in 3.3 into $r_{iat}(p)'\beta$.

$$r_{iat}(p) = \beta_1 1(post)_t 1(age6+)_a + \beta_2 X_{iat} + \gamma_a + \delta_t + \omega_p + \epsilon_{iat} \tag{5}$$

Where the main difference from 3.3 above is: ω_p, the assumed functional form of the baseline hazard. We specify this as duration month dummies [7]. Doing so allows for a flexible baseline hazard with an separate baseline impact of exit probabilities in each duration month. The coefficient β_1 in can be interpreted as a hazard ratio by employing the natural exponential function: e^{β}. Finally, while our main specification relies on hazard models, we also show results from linear probability models on the likelihood of exit in a given period. While our main specification is cox proportional hazard models, generally functional form does not matter with the cox proportional hazard and discrete time hazard giving nearly identical results.

[7] We pooled episodes greater than 95 months together. Data in this range are sparse with little variation in exits. Pooling helps the maximum likelihood estimation converge more quickly with no impact on the estimates

3.2 Results: Time to Permanency

We begin by estimating the impact of Northstar on the probability of exit from foster care into various types of permanency. By increasing payments in adoption and TPLPC to equal those in foster care for foster children age 6 or older in the post-2014 Northstar reform, the reform increased the monthly probability of exiting foster care into adoption and into TPLPC. Estimates from the Cox model support this conclusion (Table 2: Panels A & B: Column 1) and are robust when controls for child demographics, reasons for removal, and child welfare history are added (Column 2). The presented hazard ratios are interpreted as a relative increase in the rate of exit beyond the comparison group. For adoptions we estimate that there is a proportional increase of 73% $((1.73 - 1) * 100)$ for older relative to younger children after the policy. For TPLPC the proportional increase is 68%.

To assess if our results are robust to the assumptions of a hazard model, we estimate equation 3.3 above using a discrete-time hazard model (Columns 3-4) and a linear probability model (Columns 5-6) on the likelihood of exit. In all models, estimates imply increased probability of exit into adoption and TPLPC without controls (Columns 3 and 5) and with (Columns 4 and 6). The discrete time hazard results are directly comparable to the Cox hazard and similar in magnitude. The linear probability models are expressed percentage point effects on the monthly probability of a given exit type, so estimates suggest about a third of a percentage point increase monthly.

Conceptually, an increase in probability of exit to adoption or TPLPC can come at the expense of a combination of additional months of foster care and reduced probability of exit into other dispositions, such as family reunification or aging out of foster care. We see evidence of an increased probability of exit into any kind of permanency (Panel C).

To provide falsification evidence on whether the post-2014 growth in difference estimated reflects a jump in the difference after 2014 or whether they reflect different trends between the older and younger groups *before* the Northstar reform, we estimate calendar-year specific "effects" of the Northstar reform using the Cox specification from Column 2 of Table 2 in an event study type design. We normalize against the difference in outcomes between older and younger foster children's cases in 2014. In years leading up to 2014, the difference in outcomes between groups was similar to 2014, providing evidence against differential pre-trends across the different types of exits (Figure 1). We do not see much "effect" before the payment equalization actually happened. In fact, substantial jumps in the older-younger difference in exit to adoption and TPLPC start promptly after 2014, and (largely) remain at a higher level throughout the sample.

Next, we assess robustness to different choices of comparison and treatment age. We tighten our focus to cases of children closer to the sixth birthday age threshold, where post-2014 policy embeds a discontinuous shift in payments, to make the treated and comparison groups more homogeneous at the expense of losing the precision that sample size brings (Table 3). Column 1 reproduces the Table 2 Cox with controls estimates for comparison.

Column 2 estimates the same model but excludes cases of foster children in their first two years of life and children ages 10 and up. Columns 3 and 4 restrict attention only to cases of foster children within 3 and 2 years of the age threshold, respectively. As the age bandwidth and sample sizes narrow, the estimated effect on exit to adoption and TPLPC diminishes but remain statistically and substantively significant. Overall, focusing on the tightest age band (4 to 7) implies a 52% relative increase in the rates of adoption and a 28% increase in TPLPCs (relative to the comparison group). The smaller estimates could be due to reducing bias through more-homogeneous comparisons or by excluding populations with larger treatment effects (perhaps older children), in which case focusing on children 4-7 picks up a different local average treatment effect. Regardless, we consider this evidence our results are robust to reasonable changes in the specification.

A potential concern is that a move to permanency induced by a monetary stipend may be less stable than the pre-treatment permanency arrangements. We investigate this by restructuring the data such that an episode now begins after an exit from foster care into some permanency arrangement (TPLPC, reunification, or other types of arrangements). Treatment is assigned based on having been 6 or older and having exited after 2015. We then estimate our models with the outcome being the likelihood of exit from permanency into foster care. Results in table 4 show that there is no statistically significant association between receiving additional payments under Northstar and reentry into foster care. We take this as implying that the increased rates of exit do not come at the cost of entry into more fragile arrangements. This leads into an investigation of long term child human capital outcomes.

3.3 Estimation: medium-run, broader child outcomes

To look at effects on a broader set of medium-term child outcomes beyond the foster care window, such as standardized academic achievement test scores in schools, we modify the estimation strategy described in equation 3.1 above. To characterize each child's exposure to policy, we want to measure each foster child's potential exposure to Northstar during their CPE. However, we do not want simply to use the observed time during each foster spell spent in the Northstar regime because time in foster care and type of exit is, as we've documented above, endogenous to the policy.

Instead we draw from the literature on simulated policy instruments, to simulate the likelihood of Northstar exposure based on case characteristics fixed at the time of foster episode start and observable consistently across the policy regimes (Currie and Gruber (1996)). Here, the variation in treatment exposure fundamentally comes from the age of child at the episode start, the calendar month-year of the episode's start relative to Northstar's start in January 2015, along with the method used to predict expected foster episode length. Our simulated policy exposure variable will capture the proportion of time that, at the start of their foster episode, each child is predicted to be both over the age of 6 and in the post-Northstar policy period. Specifically let L_i be expected CPE length for CPE i, a_{0i} be i's age at CPE start, and

t_{0i} be an index value for the calendar year-month of CPE start. Expected exposure to treatment comes from being in the post period and over the age of 6 during the expected CPE. For each foster care episode i, the predicted share of predict spell length L_i spent post policy change is:

$$Sharepost_i = \max\{[L_i - (t_{Jan2015} - t_{0i})]/L_i, 0\}$$

If the CPE starts in or after January 2015 ($t_{Jan2015} \leq t_{0i}$), then $Sharepost =$ 1; the whole episode will occur post policy change. If the CPE starts L or more months before January 2015 ($L_i \geq> t_{Jan2015} - t_{0i}$), then $Sharepost = 0$; none of the predicted CPE will be exposed to Northstar. If it starts in the $L - 1$ months before January 2015, then $Postshare$ is the share of months of the predicted CPE during or after January 2015 and will be between 0 and 1.

Similarly, we can define the proportion of the CPE that the child is expected to be age 6 or older. If $a_{0i} \geq 6$ than the share of time in the treated group is 1. Otherwise, $Shareold = max\{[L_i - (a_6 - a_{0i})]/L_i, 0\}$. This is analogous to $Sharepost$ above.

Finally, we want a treatment variable to capture the share of predicted CPE months where the foster child would be both in the post-policy change period and 6 or older. Because both age and calendar year increase in time, we define $W = \max\{(t_{jan2015} - t_{0i}), (a_6 - a_{0i})\}$, which captures the number of months from the start of the CPE until the child is both six or older and post-policy change. Then, we construct the treatment variable such that if $t_{0i} \geq t_{jan2015}$ and $a_{i0} \geq a_6$ the share of the CPE treated is equal to 1. Otherwise,

$$Sharetreated = \max\{[L - W)]/L, 0\}$$

. If, at CPE start, it takes longer than the expected length of the CPE both for the child to be 6 and to enter the post-2015 period, then $Sharetreated = 0$. Otherwise, $Sharetreated$ is the fraction of months out of the predicted CPE length where the child is both in the post period and 6 or older.

We predict the expected length of the CPE using pre-period data. In the simplest approach, we simply set L_i to the overall average pre-period CPE length of approximately 7.4 months. In our preferred approach, we use OLS to predict, regressing CPE length on a variety of CPE characteristics fixed at CPE start.[8] We use the coefficients from this regression to predict CPE length for every case in the sample. We also, test the sensitivity of our results to assuming $L_i = 0$. This approach only counts as treated those who are fully treated for the entire CPE, so if the CPE starts on after Janurary 2015, it is in the post period and only those who start then with age at CPE start 6 or older are in the treated group. Finally, we use data for predict CPE length using regression data but using data throughout the entire sample

[8]We regress CPE length in months on age in months, race dummies (Black, White, Asian, American Indian, Pacific Islander, and unknown), a Hispanic ethnicity dummy, reason for removal dummies (physical abuse, neglect, care taker drug use, child behavior, or other), a female indicator, and a the child's number of prior CPEs.

rather than just in the pre-policy change period. We find that our results are qualitatively similar regardless of the method we use.

With a predicted value for L_i is is now possible to estimate the difference-in-differences effect as a linear model (an adaptation of :

$$Y_{iat} = \alpha + \beta_1 Sharetreated_{a_0 t_0} + \beta_2 X_{iat} + \gamma_{a_0} + \delta_{t_0} + \epsilon_{iat} \tag{6}$$

Where a and t now index year-month of CPE start and age at CPE start. Note we can control separately for $Shareold$ and $Sharepost$, or they are subsumed by time and age fixed effects δ_{t_0} and γ_{a_0}. Because we are investigating long term outcomes we define Y_{iat} for separate outcomes based on years since CPE start. For example, in the case of school attendance we could look at attendance the year of CPE start, two years after CPE start, three years after CPE. CPE start one Since the average time to adoption or TPLPC is xxxx, we consider the year of CPE start and one year after CPE start to be intermediate outcome: where many children who end up being adopted could still be in a CPE. Outcomes observed two to five years will, on the other hand, reflect impacts post CPE exit and after their child has been in their permanent home for some time.

When looking at test score outcomes, there are additional constraints on the timing of when we can look at outcomes. Standardized testing starts at the end of third grade (typically, age 8-9 years). Because children in our control group (younger than 6) don't start third large grade until 3 years after the CPE starts , we can only look at these[9]. Likewise, going more than three years out is also pushes our data for two reasons. First, the fewer children we have the farther out we go, and we already begin to lose our sample size as the education data only goes through 2019 (we lose anyone whose CPE begins after 2015 when looking at four years out, which also gives us only one year of post period). Further, our identification strategy relies on having young and old children in the both the pre- and post-2014 periods. For the post period we will observe the most young kids whose CPE starts in the post period (2015) three years out. As we go more years out we lose additional young children in the post period, because their CPEs cannot have started yet. For all these reasons we prefer to begin looking at three years out for our long term, academic outcomes. We then explore the timing in the next section.

3.4 Results: Child Outcomes

We begin by estimating impacts on school and education related outcomes. Results on test scores are shown in Figure 2. Panel A looks at the difference-in-differences effect on each child's average student achievement z-score (averaging any math and reading score available) three years after CPE start as a function of the year of CPE start. The year immediately prior to Northstar adoption (2014) is normalized to a zero effect, expressing the baseline

[9]there are some children with test scores 2 years after the CPE starts but this is a small group. We try averaging 2-3 years together in some specifications to leverage this data

difference between older and younger kids z-scores three years after their CPE starts in 2014. For CPE starting in 2012 and 2013, the pattern of later difference student achievement differences between age groups is very similar to that in 2014, yielding null estimated difference-in-difference placebo "effects" in these years. For CPE starting in Northstar's first two years (2015 and 2016), a different pattern emerges, with the (older-younger) relative difference shifting higher, yielding positive difference-in-difference effect estimates on student achievement. This pattern is especially prominent for math (Panel B).

The results are largely robust to a range of different specifications. We show results separately for reading and math scores, as well as averaging over years 2-3 rather than just looking at 3 years from ÇPE start. We test robustness to using different ways of simulating treatment through predicting CPE length: predicting off of pre-period data, simply imputing the average (pre-period) CPE length, using the interaction of CPE starts in the post period with older than 6 at CPE start (post*old), and predicting CPE length using the full sample rather than just the pre-period. Since many children in the child-welfare system are only removed from their homes for a brief time and are reunified quickly we also test the robustness by limiting the sample to those who are most likely to not be reunified. While we do not want to condition on post treatment outcomes (such as reunification), we instead predict the likelihood of being reunified using the pre-period data and our main covariates: we then remove children from the sample who are predicted as having aigh likelihood of being reunified from the sample (greater than 80 %) as a robustness test. We run every combination of the above, both with and without time variate covariates for a total of 64 specifications. These are presented as a specification curve in Figure **??**.

Some interesting patterns reveal themselves in **??**. First, the vast majority of the specifications are statistically significant, and every one of them is positive. Impacts on math Z-scores are higher than reading: this is consistent with the education literature that tends to more consistently find impacts of policy interventions on math than reading. Additionally, we tend to find larger effects when CPE length is predicted using the entire sample rather than just the pre-period (though using the pre-period for prediction remains our preferred specification. Finally, using the sample that excludes those CPEs likely to reunify results in somewhat larger impacts. This is expected as those who were removed are the least likely to be impacted by the policy.

It is also of key importance to test the parallel trends assumption. As we did with timeliness to permanency, we estimate an event study, modifying 6 by replacing $sharetreated_{a_0 t_0}$ with a series of interactions between year of CPE start and *shareold*. This estimates, for each year, the proportion of the CPE the child is predicted as being over 5 (thus, the treatment variable ranges from zero to one) [10]. Here all years post 2014 reflects a CPE that entirely occured within the post period. We show these results for average test scores,

[10]we get a similar pattern of coefficients when we estimate effects simply interacting year of cpe start with child's age at cpe start

and then separately by reading and math in figure 2. Comfortingly, we see no signs of a pre-trend across these figures. There is an increase in test scores between 2011 and 2012 before remaining flat for two years. We do not see an impact of the policy until 2016, though there is a slight increase in math scores in 2015. This could be explained by it taking time for DHS to adapt to the implementation of the policy. This does somewhat match the trends for TPLPC shown in figure Figure 1) which show a relatively small increase on exit rates in 2015 that become larger by 2016 (though a similar trend is not followed by adoptions). The pre-trends are flatter and the treatment effect larger for Math. This is also comforting as increased achievement in math is driving our main findings.

3.5 Mechanisms and Short/Long run Dynamics

Next we look more at other outcomes as well as differences in timing. While we cannot look at academic outcomes before 3 years out: we can look school attendance and stability which we have data on for every year. Results 1, 2, 3, and 4 years after the CPE starts are shown in table 5. Interestingly, we see an initial decline in attendance which fades out and then becomes a positive attendance effect as time since CPE start increases. This is possibly due to disruptions from the transition that involves time spent in court. The policy also reduces the likelihood that children change schools in each of the four years after their CPE starts. We also show an additional outcome in this table: an indicator for whether the child receives mental health services.

This outcome is somewhat difficult to interpret because it is a function of both underlying child mental health and access. Regardless, we see an initial decline in service use through this channel. If we assume that DHS policy of providing care to children who needed it was effective: then those who are moved more quickly out of foster care reflect gains in mental health. However, we can't rule out that other factors disrupted these services (such as the child living in a new area, or having less monitoring due to no longer being a ward of the state).

One challenge with all results in table 5 is that since our data only goes through 2019, we mechanistically lose observations as we look further out. Therefore it becomes unclear if changing patterns of effects are due to dynamics over time, a changing sample, or diminished statistical power. This is another reason why we focus on 3 years post CPE in our main results.

3.6 Heterogeneity

We now estimate heterogeneity on educational outcomes by age and other characteristics. As in the first stage results above, because age is essential to the identification strategy we instead look at "bandwidths" beginning for those ages closest around the 6 year cutoff to qualify for enhanced payments under Northstar through to the entire sample. These results are shown in Table 6. Overall, we see similar impacts on test scores regardless of the age restrictions.

Next we estimate heterogeneity of effects on academic achievement by different characteristics of the CPE (child gender, ethnicity, race, and reason for removal) in Table 8. Several interesting patterns emerge. Impacts on white children and Native American children are similar to the full sample effects. Impacts on Black and Hispanic children appear smaller. Benefits for are largest for children removed for neglect and physical abuse (though this is not significant statistically).

4 Conclusion

In this paper we investigates how adoption subsidies impact child well being. We argue that the our policy experiment identifies two first order effects for older relative to younger children: an equalization of payments between foster care stipend and adoption, as well as additional cash benefits flowing to older children. Our evaluation found large and substantial impacts on both the timeliness to adoption and guardianship (TPLPC) care. We found no suggestion that TPLPC placements were more likely to be dissolved, causing the child to re-enter foster care. This implies a direct benefit to the state in investing in providing stipends to perspective guardians and adoptive parents. Beyond that, we also document substantial benefits to children. Instead of the marginal adopted child, incentivized by additional stipend, entering into a less stable environment: our evidence suggests large increases in child stability and achievement three years after the CPE end.

References

Argys, L. and B. Duncan (2013). Economic incentives and foster child adoption. *Demography 50*(3), 933–954.

Baker, A., D. F. Larcker, and C. C. Wang (2021). How much should we trust staggered difference-in-differences estimates? *Available at SSRN 3794018*.

Bald, A., E. Chyn, J. S. Hastings, and M. Machelett (2019). The causal impact of removing children from abusive and neglectful homes. Technical report, National Bureau of Economic Research.

Bowles, S. (2016). *The moral economy: Why good incentives are no substitute for good citizens*. Yale University Press.

Buckles, K. S. (2013). Adoption subsidies and placement outcomes for children in foster care. *Journal of Human Resources 48*(3), 596–627.

Cuddeback, G. S. (2004). Kinship family foster care: A methodological and substantive synthesis of research. *Children and youth services review 26*(7), 623–639.

Currie, J. and J. Gruber (1996). Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of political Economy 104*(6), 1263–1296.

Doyle, J. J. (2007). Can't buy me love? subsidizing the care of related children. *Journal of Public Economics 91*(1-2), 281–304.

Doyle, J. J. and H. E. Peters (2007). The market for foster care: an empirical study of the impact of foster care subsidies. *Review of Economics of the Household 5*(4), 329.

Doyle Jr, J. J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *American Economic Review 97*(5), 1583–1610.

Doyle Jr, J. J. (2008). Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care. *Journal of political Economy 116*(4), 746–770.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research.

of Applied Research, I. (2011). Minnesota permanency demonstration: Final evaluation. Technical report, Minnesota Department of Human Services.

Roberts, K. V. (2019). Foster care and child welfare.

Service, C. R. (2019). Youth transitioning from foster care: Background and federal programs.

Testa, M. F. and N. Rolock (1999). Professional foster care: A future worth pursuing? *Child Welfare 8*, l.

Testa, M. F. and K. S. Slack (2002). The gift of kinship foster care. *Children and Youth Services Review 24*(1-2), 79–108.

Warburton, W. P., R. N. Warburton, A. Sweetman, and C. Hertzman (2014). The impact of placing adolescent males into foster care on education, income assistance, and convictions. *Canadian Journal of Economics/Revue canadienne d'économique 47*(1), 35–69.

Wildeman, C. and N. Emanuel (2014). Cumulative risks of foster care placement by age 18 for us children, 2000–2011. *PloS one 9*(3), e92785.

5 Tables and Figures

Table 1: Summary Statistics of CPE

Variable	Mean	Standard Deviation
Panel A: CPE level Characteristics		
Number of CPEs	54,577	NA
Number of Children	43,633	NA
Average CPE length.	11.4	11.6
Age in Years at CPE Start	7.93	5.90
Average of CPEs per child	1.37	0.79
Fraction White	0.57	0.49
Fraction African American	0.30	0.43
Fraction American Indian	0.25	0.43
Hispanic Ethnicity	0.10	0.30
Removed for Neglect	0.26	0.44
Physical Abuse	0.10	0.31
Caretaker Drug Abuse	0.24	0.43
Child Behavior	0.19	0.40
Reunified	0.58	0.49
Adopted	0.12	0.32
Kinship / TPLPC	0.08	0.28
Age out of Foster Care	0.02	0.16
Panel B: CPE -Month Characteristics		
of CPE-Month observations	699,413	NA
Fraction Post Northstar	0.70	0.46
6 or Older	0.59	0.49
prob exit to reunification in month n	0.0450	0.207
prob exit to adoption in month n	0.009	0.095
prob exit to Kinship/TPLPC	0.006	0.080
Panel C: Linked to Long term Outcomes (3 years after CPE start)		
# of children observed in school	13552	NA
# of re-entries to foster care	255	NA
# of children w/ long term test scores	6906	NA
Math Z-score	-0.81	0.96
Reading Z-score	-0.70	0.98
Average Z-score	-0.77	0.91
% School year attendance	0.84	0.17
schools attended (school stability)	1.56	0.84
Use mental health services	0.57	0.23

Table 2: Effects of Northstar on Foster Care Exits

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Exit to Adoption			
(Age 6+) x (Post 2014)	1.78***	1.73***	1.79***	1.71***	0.36***	0.36***
	(0.11)	(0.11)	(0.11)	(0.11)	(0.05)	(0.05)
			Panel B: Exit to Kindship / TPLPC			
(Age 6+) x (Post 2014)	1.83***	1.68***	1.67 ***	1.57 ***	0.34***	0.30***
	(0.11)	(0.11)	(0.11)	(0.10)	(0.05)	(0.05)
			Panel C: Exit to Permanency			
(Age 6+) x (Post 2014)	1.46 ***	1.37 ***	1.37***	1.30 ***	0.73 ***	0.70 ***
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Model	Cox	Cox	Discrete	Discrete	LPM	LPM
Controls	No	Yes	No	Yes	No	Yes
# of Foster care spells	54,577	54,571	54,577	54,571	54,571	54,571
Observations	699,413	699,375	699,413	699,375	699,413	699,375

Notes: An observation is a year-month that a child is observed in a foster care spell. Each Column and Row are from a separate DD regression on the interaction between being age 6+ in the post Northstar period (2015+) with age and year-month fixed effects on different types of exits from foster care. Columns 1-2 show results from cox proportional hazard models. Columns 3-4 show results from discrete time hazard models. Columns 5-6 show results from a linear probability model. For the hazard models we report relative hazard ratios and their standard errors. Hazards above (below) 1 reflect a proportionate increase (decrease) of the treated group relative to the comparison group. For the linear probability model we multiply the coefficients by 100 so they reflect a percentage point likelihood of exit in a given year-month. For models with controls (the even columns), controls include:race (white, African-American/Black, Native American, Asian, Pacific Islander, Unknown, and other), Hispanic Ethnicity, reason for removal (neglect, physical abuse, care taker drug use, behavioral problems, and other), gender, and a child's total number of foster care placements.

Figure 1: Event Study Cox Hazard

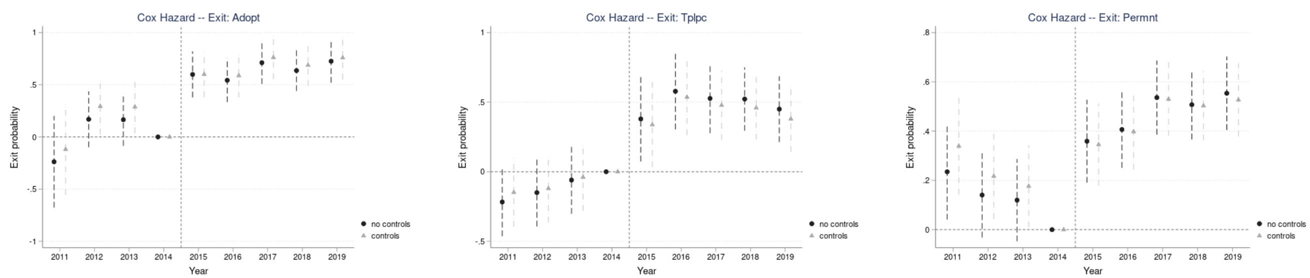


Figure 1 shows hazard ratios from estimating the event study version of 3.3. Each coefficient estimate represents the proportional increase in exits of older (6+) children relative to younger children, in a given year: normalized so that 2014=1.

Table 3: Exits from foster care: varying age bandwidth

	(1)	(2)	(3)	(4)
	Panel A: Exit to Adoption			
	All Ages	Ages 2-9	Ages 3-8	Ages 4-7
Simulated NS exposure	1.73***	1.43 ***	1.44 ***	1.52 ***
	(0.11)	(0.12)	(0.13)	(0.17)
	Panel B: Exit to TPLPC			
(Age 6+) x (Post 2014)	1.68***	1.20 ***	1.19 ***	1.28 ***
	(0.11)	(0.10)	(0.11)	(0.15)
	Panel C: Exit to Any Permanency			
(Age 6+) x (Post 2014)	1.46 ***	1.18 ***	1.22 ***	1.32 ***
	(0.06)	(0.06)	(0.08)	(0.08)
model	cox	cox	cox	cox
controls	No	No	No	No
# of Foster care spells	54,577	24,812	18,742	13,582
Observations	699,413	284,601	195,376	150,845

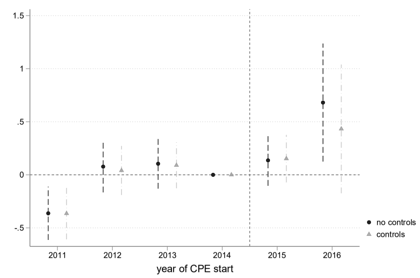
Notes: An observation is a year-month that a child is observed in a foster care spell. Each Column and Row are from a separate DD regression on the interaction between being age 6+ in the post Northstar period (2015+) with age and year-month fixed effects on different types of exits from foster care. Columns 1-2 show results from cox proportional hazard models. Columns 3-4 show results from discrete time hazard models. Columns 5-6 show results from a linear probability model. For the hazard models we report relative hazard ratios and their standard errors. Hazards above (below) 1 reflect a proportionate increase (decrease) of the treated group relative to the comparison group. For the linear probability model we multiply the coefficients by 100 so they reflect a percentage point likelihood of exit in a given year-month. For models with controls (the even columns), controls include:race (white, African-American/Black, Native American, Asian, Pacific Islander, Unknown, and other), Hispanic Ethnicity, reason for removal (neglect, physical abuse, care taker drug use, behavioral problems, and other), gender, and a child's total number of foster care

Table 4: Effect of Northstar on Re-entry into Foster care after Adoption or TPLPC

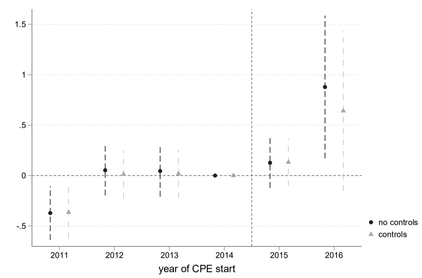
	(1)	(2)
Panel A: Re-entry from Adoption		
(Age 6+) x (Post 2014)	1.63	1.21
	(1.40)	(1.09)
of re-entries	19	19
Panel B: Re-entry from Kinship / TPLPC		
(Age 6+) x (Post 2014)	0.76	0.77
	(0.21)	(0.22)
of re-entries	236	236
Panel C: Re-entry from any Permanency		
(Age 6+) x (Post 2014)	1.03	0.97
	(0.26)	(0.25)
Model	Cox	Cox
Controls	No	Yes

Notes: An observation is a year-month that a child is observed in each type of permanency arrangement . Each Column and Row are from a separate DD regression on the interaction between being age 6+ in the post Northstar period (2015+) with age and year-month fixed effects. ALl models are estimated with cox proportional hazards. Hazards above (below) 1 reflect a proportionate increase (decrease) of the treated group relative to the comparison group. Controls include:race (white, African-American/Black, Native American, Asian, Pacific Islander, Unknown, and other), Hispanic Ethnicity, reason for removal (neglect, physical abuse, care taker drug use, behavioral problems, and other), gender, and a child's total number of foster care placements.

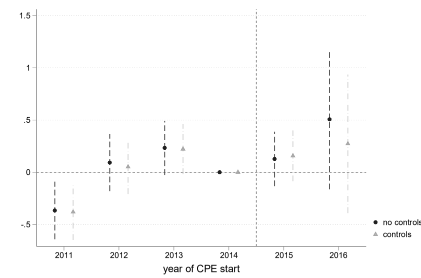
Figure 2: Event Study on Standardized Test Z-scores: 3 years out



(a) Average Z-score



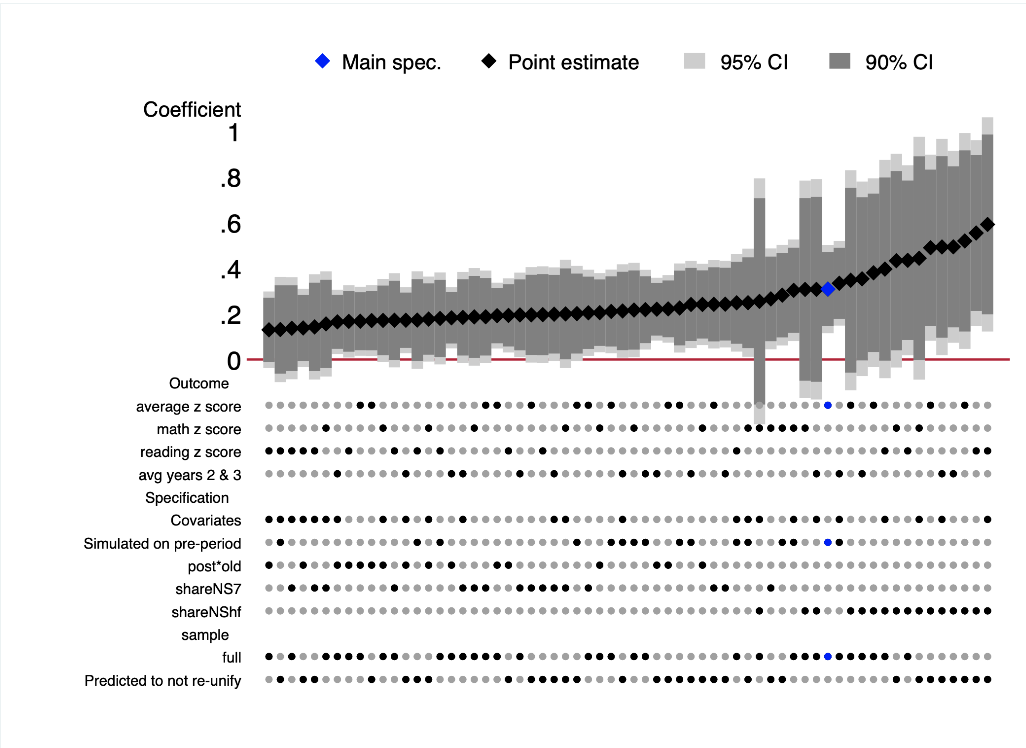
(b) Math



(c) Reading

This figure shows z-scores interacted with a continuous variable for share of time in the predicted CPE. We predicted CPE length using OLS on the pre-period data to regress length of CPE on a full set of covariates. The graph is normalized so that 2014=0.

Figure 3: Specification Curve



This figure shows coefficient estimates from a wide range of specification tests. Each test regresses (some form) of a test score measure on various difference-in-difference models.

Table 5: The Impact of Northstar on ancillary outcomes

Years after CPE start:	(1 year)	(2 years)	(3 years)	(4 years)
	Panel A: Received mental health service			
Simulated policy exposure	- 2.7 **	-1.6 **	-2.0 **	-6.7 **
	(0.80)	(0.80)	(0.90)	(3.3)
Pre-2015 mean	10.0	8.0	6.0	5.0
of CPEs	22331	17652	10637	9671
.				
	Panel B: School Attendance			
Simulated policy exposure	-1.2 **	-0.4	0.9 *	3.7 *
	(0.50)	(0.50)	(0.50)	(2.10)
Pre-2015 mean	88	88	89	90
of CPES	22331	17652	10637	9671
	Panel C: # of schools per year			
Simulated policy exposure	-0.08 **	-0.06 **	-0.06 *	0.34 ***
	(0.03)	(0.03)	(0.03)	(0.12)
Pre-2015 mean	1.86	1.68	1.58	1.51
of CPEs	22331	17652	10637	9671
controls	no	no	no	no

Table 6: Heterogeneity: Effects by Age

.	All Ages	2-9	3-8	4-7
	Panel A: Average Test Z-Scores			
Simulated Policy	0.31 ***	0.26 **	0.25 **	0.21 **
	(0.10)	(0.10)	(0.11)	(0.11)
Pre-2015 mean	-0.78	-0.73	-0.73	-0.69
of CPES	6906	4595	3770	2906
	Panel B: Math Z-Scores			
Simulated Policy	0.31 ***	0.29***	0.27 **	0.25 **
	(0.10)	(0.11)	(0.11)	(0.11)
Pre-2015 mean	-0.81	-0.76	-0.75	-0.71
of CPEs	6140	4535	3727	2878
	Panel C: Reading Z-scores			
Simulated Policy	0.25 **	0.21 *	0.21 *	0.16
	(0.11)	(0.11)	(0.11)	(0.12)
Pre-2015 mean	-0.72	-0.70	-0.70	-0.67
of CPEs	6158	4561	3750	2892
controls	no	no	no	no

Table 7: Heterogeneity in Effects of Northstar on Foster Care Exits: Permanency

Panel A: Heterogeneity by gender

(Age 6+) x (Post 2014)	All	Male	Female	
	1.46 ***	1.43 ***	1.49 ***	
	(0.06)	(0.09)	(0.09)	
# of CPEs	54,577	28,994	25,583	
# of observations	699,413	362,904	. 336,509	

Panel B: By Reason for Removal

	Neglect	Physical Abuse	Caretaker Drug use	
(Age 6+) x (Post 2014)	1.27 ***	1.47 ***	1.10	
	(0.09)	(0.21)	(0.09)	
# of CPEs	14,242	5,691	12,955	
# of observations	219,220	65,680	77,069	

Panel C: Heterogeneity by Race / Ethnicity

	White	Black	Native	Hispanic
(Age 6+) x (Post 2014)	1.39 ***	1.66 ***	1.57 ***	1.57 sym**
	(0.08)	(0.15)	(0.18)	(0.14)
# of CPEs	26,677	15,528	12,190	5,597
# of observations	322,234	178,947	67,849	122,185
Model	Cox	Cox	Cox	. Cox.
Controls	no	no	no	no

Table 8: Heterogeneity in Effects on Standardized Test Z-Scores, by Foster Child and Case Characteristics

.

	Panel A: Heterogeneity by gender			
	All	Male	Female	
Simulated policy exposure	0.31 ***	0.55 ***	0.11	
	(0.099)	(0.13)	(0.14)	
Pre-2015 mean	-0.78	-0.87	-0.69	
# of CPEs	6906	3507	3399	
	Panel B: By Reason for Removal			
	Neglect	Physical Abuse	Caretaker Drug use	
Simulated policy exposure	0.39 **	0.35	0.10	
	(0.16)	(0.27)	(0.19)	
Pre-2015 mean	-0.79	-0.95	-0.57	
# of CPEs	2266	939	1563	
	Panel C: Heterogeneity by Race / Ethnicity			
	White	Black	Native	Hispanic
Simulated policy exposure	0.38 **	0.16	0.39 **	0.09
	(0.15)	(0.18)	(0.18)	(0.26)
Pre-2015 mean	-0.6	-1.11	-0.83	-0.87
# of CPEs	3595	1785	1741	702
controls	no	no	no	no