**Abstract:**

This paper explores the effects of attrition on student development in early education. Utilizing data from Tennessee's Project STAR experiment, this paper aims to expand upon the literature of peer effects as well as competition between public and private schools. It first reproduces the attrition profile of Project STAR first outlined in Rohlfs \& Zilora's (2014) to determine the nature of attrition in the experiment. It departs from other papers in the literature by utilizing survival analysis to determine which characteristics of students prolonged participation in the experiment. This paper then uses these findings to estimate the peer effects of attrition on students who remained in the experiment through the change in test scores of students between the 1\textsuperscript{st} and 2\textsuperscript{nd} grades. Such peer effects are subsequently decomposed using clustering to estimate the peer effects from students who may have left to private schools.

This paper aims to provide evidence that student sorting across different types of schools has educational impacts on the students they leave behind. At the same time, it analyzes attrition in Randomized Controlled Trials and may provide evidence of subsequent confounding factors or spillover effects on the broader population related to such developments.

\vspace{.25in}

\noindent \textit{JEL Classification:} I, I21, I26, H4, J13.

\noindent \textit{Keywords:} Attrition, Economics of Education, Peer Effects, Private Schools, Project STAR.

**Introduction:**

**Lit Review:**

* Include more stuff on peer effects
* Include more stuff on low income mobility
* Include more stuff on development and education
* Reframe question and context
* Actually read papers (especially ones w/ influence on data)

**Data:**

* Read other papers to more accurately summarize data
* Understand and decide on summary charts
* Format charts and footnote correctly
* Explain measures, variables, etc.

**Empirics:**

* Update empirical specifications according to notes
* Rewrite final section and peer effects section to better explain what I’m doing and why
* Read Jessica’s/TJ’s to understand everything

**Results:**

**AFT K-1**

* Specification 1 is fitted on all students in the kindergarten cohort, while specification 2 is fitted on the subset that stayed beyond 1st grade (hence they have the opportunity to switch schools or class types). Specification 1 allows us to see the total effects of attrition across all years in the experiment and specification 2 aims to reduce the effects of attrition that arose from re-randomization or school constraints after the first year of the experiment.
* Similar to the survival curves, I decline to fit this model on any students whose schools dropped out of the experiment, once again focusing on “unforced” attrition. Additionally, I drop observations where a student in the kindergarten cohort switched to a different school in 1st grade and came back to the same kindergarten school in 2nd grade. Such students would not be considered “departed” in my experiment; however their school-to-school movement may bias my results. In total, there are 7 students who fall into this category, a very small amount. I estimate that less than .15% of total observations exhibit a similar phenomenon (multiple school switches back to original school) across all years of the experiment.
* The results from specification 1 are consistent with those found in the non-parametric estimates in the previous section. Specifically, students who are not on free lunch are expected to survive $e^{0.287} \approx 1.332$ times longer than students who are on free lunch, holding all else constant. Students in rural schools are expected to survive $e^{0..341} \approx 1.406$ times longer than students in inner-city schools and similarly students in urban schools are expected to survive $e^{0.141} \approx 1.151$ times longer than students in inner-city schools, holding all else constant. With regard to ability, I find that students with kindergarten test scores between the 25th and 75th percentiles of all students in the kindergarten cohort are expected to survive $e^{0.506} \approx 1.657$ times longer than those in the 25th percentile or lower, holding all else constant. Similar results are observed for students in the 75th percentile or higher who are expected to survive $e^{0.741} \approx 2.098$ times longer than those in the 25th percentile or lower, holding all else constant.
* The results for both income and ability are also observed in specification two, although with lower effect sizes. Across both specifications, these effects are significant at the 1\% level. It is also important to note that I do not find significant effects from class size on attrition among this cohort. Diagnostic plots in the appendix show that the assumption of a Weibull distribution to model attrition is well-founded. Attrition appears to monotonically decrease as the experiment goes on and the residuals appear to converge.

**AFT 1-2**

* The below table shows regression results for the AFT model of student attrition in STAR among all students in 1st grade. This includes new students who entered in the 1st grade cohort as well as those that stayed from kindergarten. Similar to the previous section, specification 1 is fitted on all students in 1st grade, while specification 2 is fitted on those who stayed for at least 1 year.
* Among this cohort, the results are mixed. With regard to ability, the results from specification 1 are consistent with those in the previous section. Students above the 25th percentile are expected to stay in the experiment longer than those below the 25th percentile based on 1st grade test scores. These effects are insignificant among those who stayed past 2nd grade. This is potentially due to the fact that attrition in the final year of the experiment was quite small, with little variation among students that left.
* It is important to note that in specification 1, I find that students in inner-city schools are expected to survive longer than those in their suburban, rural, or urban counterparts. This is consistent with the finding that being on free lunch is no longer a significant factor in attrition, with it in specification 2 actually contributing to an increase in survival time. Class size effects are also significant in both specifications, with students in regular-aide classes more likely to survive than their small-class counterparts.
* It is most likely the case that attrition among the kindergarten students was more pronounced along income lines; however, students with low-ability were less likely to stay in the experiment, regardless of their entry cohort.

**Peer effects K-1**

**Peer effects 1-2**

**Clustering**

**Discussion:**

* Writeup, limitations, conclusions
* Policy Impact

List of running questions:

* For survival variable, do I use number of years in star? This is a composite, non time-indexed measure (i.e. it doesn’t care if the student was in star between 1980-1981, only cares aggregate number of years)
  + In principle, my functional form doesn’t change if it is time indexed or not, either way attrition is still monotonically decreasing (i.e. more for 0-1 than 1-2)
  + Also, since I measure by entry cohort, doesn’t necessarily matter for if a kid left btwn 2nd and 3rd
  + Am I still teasing out variation especially due to changes between K and 1 due to re-randomization? (by controlling for class size should be ok right)
* Right now I’m using my own variables, I think the coding of the composite duration is incorrect because I want to understand that if they switched to a small class (i.e. the parent was able to lobby, does that lessen their chance of attrition? Tells smtg about preferences)
  + Can switch to composite metric later, but I believe it is wrong
* Do I need to show hypothesis tests for randomization? I think not b/c already well established, but let me know.
* Do I need to control for exit variables (e.g. teacher gender on exit? Will this somehow confound for students that stayed the whole time?)