**Bifulco, R., Rubenstein, R., & Sohn, H. (2017). Using Synthetic Controls to Evaluate the Effect of Unique Interventions: The Case of Say Yes to Education. *Evaluation Review, 41*(6), 593-619. https://journals.sagepub.com/doi/abs/10.1177/0193841X17742233**

On variables used to find weights:

Researchers must also choose the linear combinations of pretreatment outcome observations the synthetic control should seek to match. One natural choice is the value of the outcome variable for all the available pretreatment periods. In practice, including all available pretreatment outcomes measures does not necessarily result in a closely matched synthetic control. For instance, some of the pretreatment outcome values might be disproportionately influenced by transitory shocks, and it is not clear that closely matching those values will improve estimates of the counterfactual outcomes for the treated district during the posttreatment period. The existing literature provides little guidance about how to choose the set of linear combinations of pretreatment outcome observation periods.

To explore the sensitivity of effect estimates to the choice of pretreatment outcome observations, we construct synthetic controls using six different sets of pretreatment years, which are detailed in Table 1. To allow for assigning weights to districts that are similar to SCSD in terms of socioeconomic characteristics, we also include in all of our specifications the average shares of free-lunch eligible, African American, and Hispanic students over the entire pretreatment period, which for the enrollment analysis ranges from 1998–1999 to 2007–2008 and for the graduation analysis covers the 1998 through 2004 cohorts. The inclusion of covariates (e.g., the average share of free-lunch eligible) when estimating the weights reduces the likelihood that some districts quite dissimilar demographically to Syracuse could contribute to the counterfactual outcome.

**Table 1.** Alternative Specifications of Pretreatment Years.

| Specification | Description |
|---|---|
| 1 | First and last year of pretreatment periods |
| 2 | First, middle, and last year of pretreatment periods |
| 3 | Middle and last year of pretreatment periods |
| 4 | Last pretreatment year and the average of outcomes in all other pretreatment years |
| 5 | Each pretreatment year |
| 6 | Each year from the middle to the end of the pretreatment periods |

Weights used under different specifications of pretreatment years (enrollment outcome):

**Table 2.** Assignment of Weights (Enrollment Analysis).

| | Assigned Weights | | | | | |
|---|---|---|---|---|---|---|
| District Name | Specif. 1 | Specif. 2 | Specif. 3 | Specif. 4 | Specif. 5 | Specif. 6 |
| Panel A: Restricted donor pool | | | | | | |
| Albany CSD | .000 | .021 | .000 | .000 | .005 | .000 |
| Brentwood UFSD | .000 | .000 | .000 | .000 | .116 | .129 |
| Buffalo CSD | .000 | .078 | .022 | .034 | .029 | .000 |
| Niagara Falls CSD | .484 | .288 | .498 | .499 | .411 | .404 |
| Rochester CSD | .502 | .406 | .479 | .467 | .438 | .467 |
| Utica CSD | .014 | .207 | .001 | .000 | .000 | .000 |
| Panel B: Comprehensive donor pool | | | | | | |
| Albany CSD | .000 | .000 | .000 | .000 | .005 | .000 |
| Brentwood UFSD | .000 | .000 | .000 | .000 | .116 | .134 |
| Buffalo CSD | .117 | .174 | .091 | .065 | .029 | .004 |
| Elmira CSD | .000 | .000 | .307 | .148 | .000 | .000 |
| Hopevale UFSD | .101 | .197 | .053 | .021 | .000 | .000 |
| Mount Vernon CSD | .000 | .061 | .000 | .000 | .000 | .000 |
| Niagara Falls CSD | .248 | .000 | .124 | .324 | .411 | .401 |
| Rochester CSD | .386 | .341 | .425 | .442 | .438 | .461 |
| Smithtown CSD | .149 | .156 | .000 | .000 | .000 | .000 |
| Utica CSD | .000 | .069 | .000 | .000 | .000 | .000 |

*Note*. Specif., CSD, and UFSD denote "specification," "city school district," and "union free school district," respectively. Districts that do not appear in the table do not receive positive weights equal to or greater than 0.001 in any of the specifications. Restricted donor pool includes 22 districts, whereas comprehensive donor pool includes 275 districts.

## Results: enrollment outcome

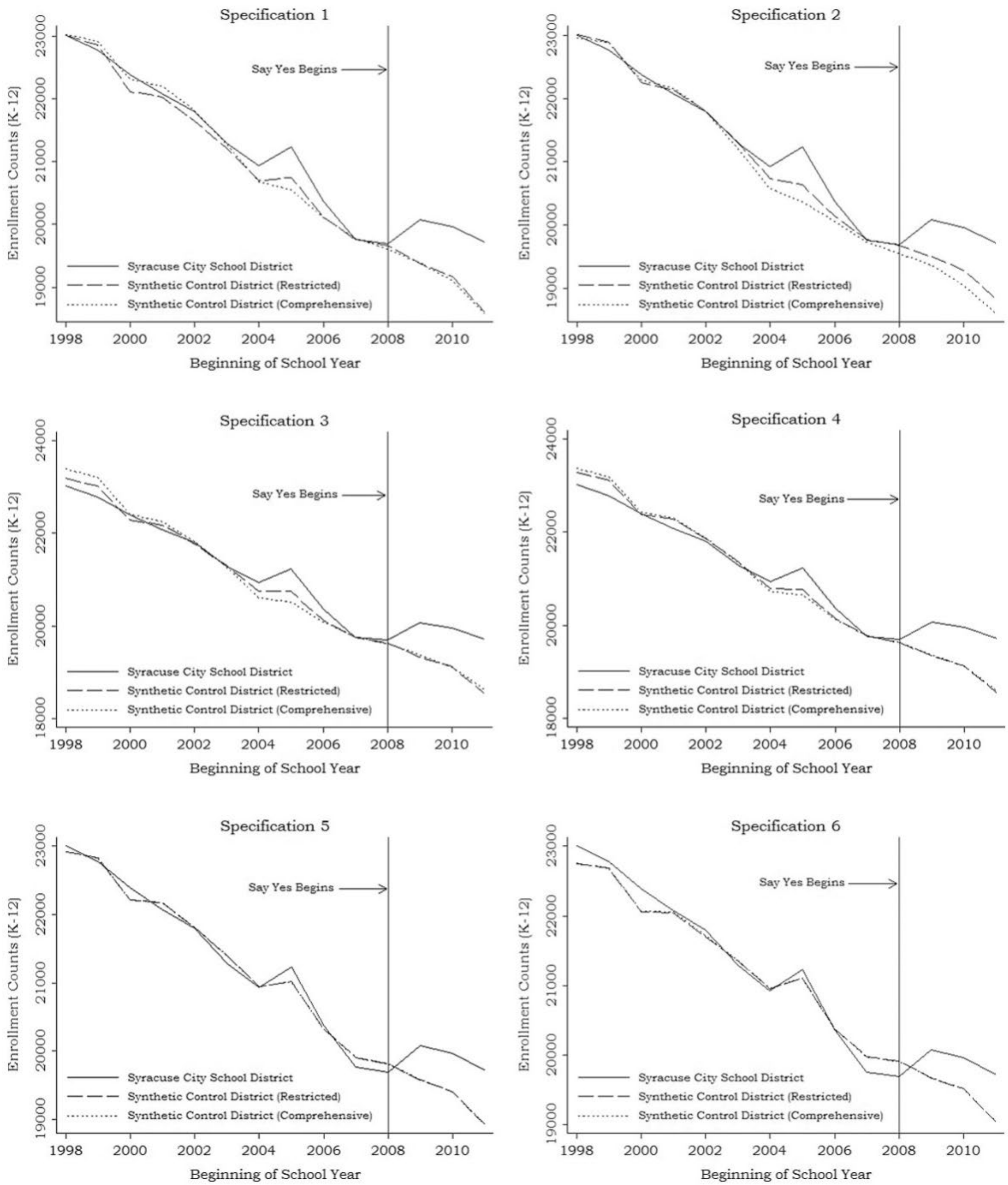


**Figure 2.** Trends in enrollment by model specifications. See Table 1 for description of pretreatment years included in each specification.

On inference:

### *Inference Procedure*

Because effect estimates are based on comparison of a single treated unit with a synthetic control, an asymptotic approximation is not well suited for assessing the uncertainty in the estimates. Abadie et al. (2010) propose a permutation test to assess the likelihood that the effect size found in the treated unit is sufficiently unlikely to occur by chance. Specifically, they recommend constructing a synthetic control for each unit in the sample, using the remaining districts as the donor pool for each other district. The true treatment effect in each of these untreated units is presumably zero. By comparing the "effect estimate" in each of the untreated units to that found in the treated unit, one can assess how likely it is to obtain an effect estimate that large by chance.

A stylized example helps to illustrate the procedure. Assume we have 1 treated unit and 99 untreated donor units. The treated unit receives the treatment at time $t$. We construct a synthetic control for the treated unit and find that outcome value in the treated unit increased by the amount $\Delta Y$. We then construct synthetic controls for each of the untreated donor units, assuming that these untreated units also received the treatment at time $t$, and measure the increase in the outcome value in each, relative to its synthetic control, in the posttreatment period. If we find that none of the 99 untreated district has an increase as large as or larger than $\Delta Y$ in the posttreatment period, this is analogous to a 99% confidence level the increase in the treated district was not due to chance.

More specifically, for each iteration of the synthetic control procedure, we compute the prediction errors for each year, square them (because the errors can be either positive or negative), and take the square root of the average of the squared prediction errors in the pretreatment period and the square root of the average of the squared prediction errors in the posttreatment period. These are called the pre- and posttreatment root mean squared prediction errors (RMSPEs). The ratio of the posttreatment RMSPE to the pretreatment RMSPE is used as a test statistic. The higher this test statistic, the greater the deviation in the posttreatment period between the district and its synthetic control and, therefore, the stronger the evidence of a policy impact. Calculating the percentage of test statistics from all iterations of this procedure that are as large as or larger than that obtained for Syracuse provides a $p$-value. Because we assume that the effect of Say Yes in all of the donor pool districts is zero, this $p$-value can be interpreted as the probability of obtaining effect estimates as large as that obtained for Syracuse if the true treatment effect were zero (Abadie, Dimaond, & Hainmueller, 2010; Cameron & Miller, 2015).

Estimated effects on enrollment:

**Table 3.** Estimated Effects on K–12 Enrollments, RMSPE, and *p*-value.

| Specification | Year 1 | Year 2 | Year 3 | Year 4 | RMSPE | *p*-value |
|---|---|---|---|---|---|---|
| Panel A: Restricted donor pool | | | | | | |
| Specification 1 | 30 | 704 | 789 | 1,110 | 214.16 | .087 |
| Specification 2 | 24 | 576 | 676 | 889 | 220.09 | .217 |
| Specification 3 | 69 | 739 | 840 | 1,166 | 209.04 | .044 |
| Specification 4 | 64 | 730 | 839 | 1,165 | 227.19 | .130 |
| Specification 5 | −114 | 500 | 560 | 795 | 114.27 | .304 |
| Specification 6 | −216 | 405 | 445 | 672 | 161.23 | .130 |
| Panel B: Comprehensive donor pool | | | | | | |
| Specification 1 | 87 | 693 | 859 | 1,132 | 252.35 | .076 |
| Specification 2 | 147 | 715 | 920 | 1,117 | 318.44 | .243 |
| Specification 3 | 76 | 702 | 845 | 1,075 | 323.72 | .098 |
| Specification 4 | 67 | 713 | 840 | 1,121 | 280.36 | .091 |
| Specification 5 | −114 | 500 | 560 | 795 | 114.27 | .562 |
| Specification 6 | −213 | 404 | 445 | 668 | 156.37 | .120 |

*Note*. Restricted donor pool includes 22 districts, whereas comprehensive donor pool includes 275 districts. Years 1–4 correspond to the effect estimates. *p*-value implies a probability of getting a post/pretreatment RMSPE ratio as large as the post/pretreatment RMSPE ratio of Syracuse if one assigns the treatment at random in the data. Specifications are the same as in Table 1. Pretreatment period includes years 1998–2007. All models are run with percent Black, percent Hispanic, and percent free lunch eligible as covariates. RMSPE = root mean squared prediction error.

Weights used under different specifications of pretreatment years (graduation outcome):

**Table 5.** Assignment of Weights (Cohort Graduation Rate Analysis).

| District Name | Assigned Weights | | | | | |
|---|---|---|---|---|---|---|
| | Specif. 1 | Specif. 2 | Specif. 3 | Specif. 4 | Specif. 5 | Specif. 6 |
| Panel A: Restricted donor pool | | | | | | |
| Albany CSD | .064 | .000 | .000 | .000 | .000 | .000 |
| Buffalo CSD | .323 | .800 | .000 | .608 | .646 | .609 |
| East Ramapo CSD | .000 | .000 | .000 | .000 | .000 | .027 |
| Hempstead UFSD | .000 | .021 | .000 | .000 | .134 | .000 |
| Hudson CSD | .000 | .000 | .002 | .000 | .000 | .001 |
| Niagara Falls CSD | .000 | .101 | .160 | .000 | .000 | .000 |
| Poughkeepsie CSD | .000 | .000 | .396 | .000 | .000 | .000 |
| Rochester CSD | .351 | .000 | .000 | .157 | .000 | .000 |
| Schenectady CSD | .261 | .078 | .442 | .232 | .219 | .363 |
| Panel B: Comprehensive donor pool | | | | | | |
| Albany CSD | .042 | .000 | .000 | 0.005 | .012 | .000 |
| Buffalo CSD | .679 | .477 | .368 | 0.683 | .000 | .428 |
| Binghamton CSD | .000 | .000 | .000 | 0.000 | .131 | .000 |
| East Ramapo Central SD | .000 | .000 | .000 | .000 | .000 | .066 |
| Elmira CSD | .190 | .000 | .000 | .193 | .000 | .000 |
| Glen Cove CSD | .000 | .000 | .000 | .000 | .052 | .000 |
| Greenburgh Central SD | .000 | .000 | .000 | .000 | .000 | .019 |
| Greenburgh Eleven UFSD | .090 | .166 | .015 | .072 | .219 | .189 |
| Hawth.-Cedar Kn. UFSD | .000 | .000 | .000 | .000 | .111 | .000 |
| Hempstead UFSD | .000 | .000 | .000 | .000 | .025 | .000 |
| Niagara Falls CSD | .000 | .168 | .000 | .000 | .000 | .000 |
| Poughkeepsie CSD | .000 | .000 | .392 | .047 | .058 | .000 |
| Rensselaer CSD | .000 | .189 | .226 | .000 | .269 | .059 |
| Rochester CSD | .000 | .000 | .000 | .000 | .002 | .000 |
| Watervliet CSD | .000 | .000 | .000 | .000 | .000 | .222 |
| Westbury UFSD | .000 | .000 | .000 | .000 | .000 | .018 |

*Note.* Specif., CSD, and UFSD denote "specification," "city school district," and "union free school district," respectively. Districts that do not appear in the table do not receive positive weights equal to or greater than 0.001 in any of the specifications. Restricted donor pool includes 22 districts, whereas comprehensive donor pool includes 236 districts.

Results: graduation rates

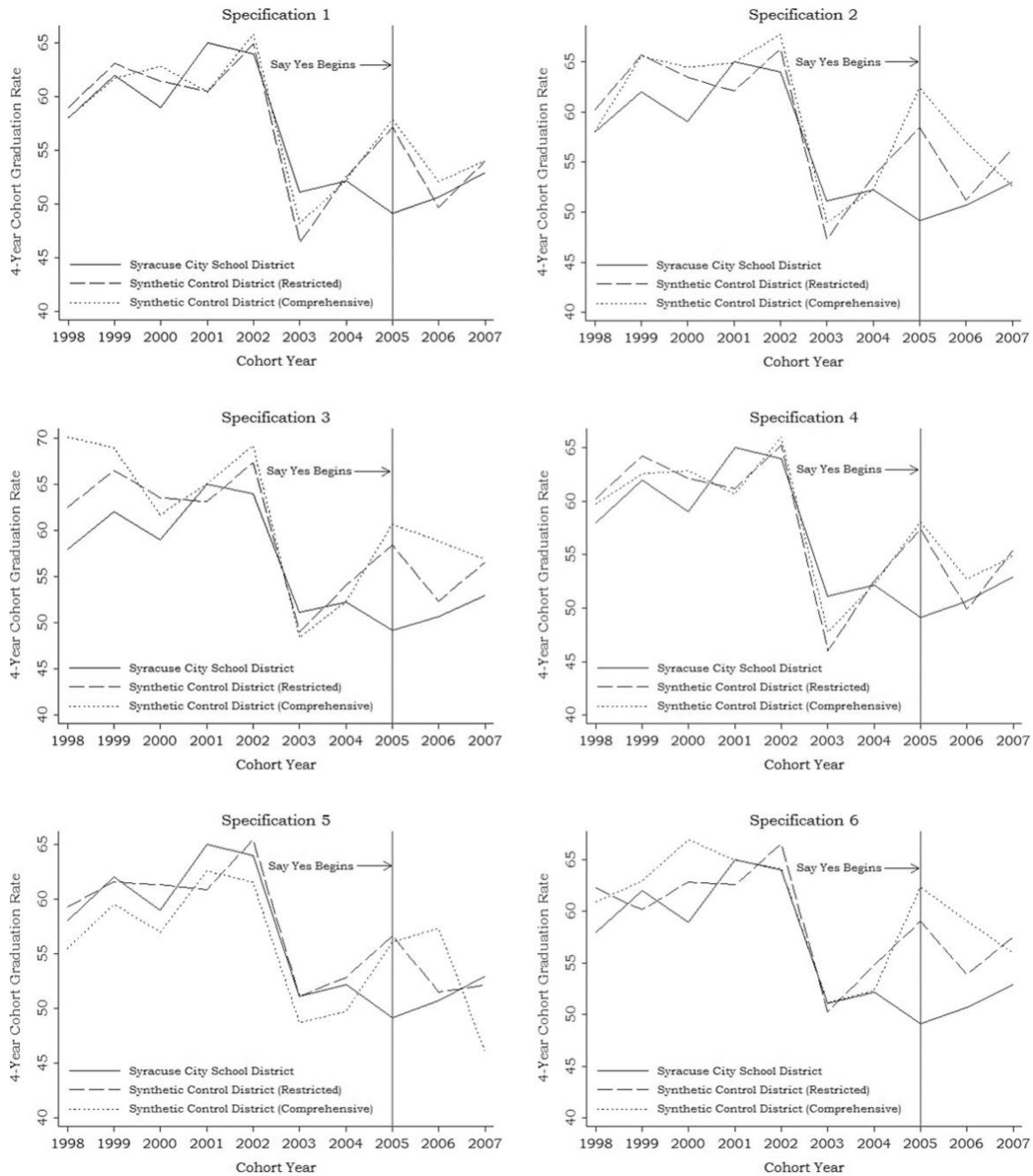


**Figure 3.** Trends in 4-year cohort graduation rate by model specifications. See Table 1 for description of pretreatment years included in each specification.

Results: graduation rates

**Table 6.** Estimated Effects on Cohort Graduation Rates, RMSPE, and *p*-value.

| Specification | Year 1 | Year 2 | Year 3 | RMSPE | *p*-value |
|---|---|---|---|---|---|
| Panel A: Restricted donor pool | | | | | |
| Specification 1 | −3.373 | −1.015 | −0.469 | 2.53 | .696 |
| Specification 2 | −9.611 | −1.763 | −3.344 | 3.14 | .435 |
| Specification 3 | −9.255 | −8.684 | −6.442 | 6.88 | .348 |
| Specification 4 | −6.076 | −0.908 | −1.911 | 2.21 | .348 |
| Specification 5 | −7.505 | −0.837 | 0.842 | 1.95 | .565 |
| Specification 6 | −8.062 | −2.268 | −3.646 | 3.26 | .174 |
| Panel B: Comprehensive donor pool | | | | | |
| Specification 1 | −8.753 | −1.421 | −1.071 | 2.52 | .101 |
| Specification 2 | −13.274 | −6.286 | 0.367 | 2.84 | .055 |
| Specification 3 | −11.527 | −8.127 | −3.870 | 5.79 | .215 |
| Specification 4 | −8.904 | −2.038 | −1.913 | 2.66 | .106 |
| Specification 5 | −6.917 | −6.630 | 6.850 | 2.48 | .557 |
| Specification 6 | −13.202 | −8.398 | −2.984 | 3.06 | .076 |

*Note*. Restricted donor pool includes 22 districts, whereas comprehensive donor pool includes 236 districts. Years 1–3 correspond to the effect estimates. *p*-value implies a probability of getting a post/pretreatment RMSPE ratio as large as the post/pretreatment RMSPE ratio of Syracuse if one assigns the treatment at random in the data. Specifications are the same as Table 1. Pretreatment period includes years 2001–2007. RMSPE = the root mean squared prediction error.

Useful concluding remarks:

## Conclusions

This article uses synthetic control methods to evaluate the effects of a place-based scholarship program and to highlight critical issues evaluators face in using the technique to construct valid counterfactual estimates. In the analysis of enrollments, where the time series of outcome measures does not appear to be strongly influenced by transitory shocks, synthetic control methods provide consistent evidence of meaningful public school enrollment increases in Syracuse relative to control districts after the start of the Say Yes to Education program. Moreover, these effect estimates did not vary substantially when choices of donor pool and pretreatment outcome years were changed, providing an important check on the robustness of the results. Because the synthetic control method disciplines the choice of comparison units used to estimate counterfactual posttreatment outcomes, we are able to derive effect estimates that are largely insensitive to the sample of units used in the estimation and remove a potentially important source of bias in effect estimates.

In contrast to enrollments, the time series of graduation rates is quite volatile. In this case, the effect estimates provided by synthetic control methods were sensitive to both the choice of pretreatment years included in the matching algorithm and the choice of donor pool districts. Given these results, and that synthetic controls can yield minimally biased effect estimates only when the number of preintervention time periods is large relative to the scale of transitory shocks (Abadie et al., 2010), it is difficult to draw conclusions about the effects of the treatment on graduation rates. It is important to note, however, that estimates derived from parametric regressions can also be sensitive to both functional form and sample choices when pretreatment outcomes trends are noisy.

More substantively, it is quite possible that meaningful effects on graduation rates in Syracuse will take time to develop. Students already at risk of dropping out when the program was announced may be unmotivated by the scholarship offer or, more importantly, may already be too far behind to graduate. As students who received additional supports in elementary and middle school begin to reach graduation age, it is possible that graduation rates may start to increase.

Finally, we find consistent evidence that, after decades of population and enrollment declines, public school enrollments in the city of Syracuse began to rebound following the start of the Say Yes to Education program. While the estimated 4-year increases of between 3% and 6% may appear modest, it