Estimating the investment effect of the Growth Catalyst Program with synthetic control methods

Balázs Muraközy
University of Liverpool Management School
March 31, 2020

This paper uses synthetic control group methods to estimate the investment effect of the Growth Catalyst, a £1.6m European funded, 8-month leadership programme for small firm entrepreneurs in the Liverpool City-Region. Given the small size of the treated group, we use synthetic control group methods to estimate investment effects both at the firm level and for the total program. Such methods can be applied in similar situations because they provide important feedback for program development by showing which firms benefitted most from the program. We find that the Growth Catalyst initiative had a significant effect on firm level investment, equivalent to about 38% increase of the fixed assets for surviving firms on average. In monetary terms, this is equivalent to £283 thousand per firm or £36.5 million for the full Growth Catalyst program between 2014 and 2018.

Keywords: SME, investment, training, synthetic control group

JEL-codes: L25, R11, C21

1 Introduction

The Liverpool City-Region suffers from several structural weaknesses including including relatively low productivity levels and low levels of entrepreneurship. Many different programs and approaches have been employed to alleviate such weaknesses. One approach is to develop managerial competencies which is a key determinant of productivity (Bloom and van Reenen, 2017).

The £1.6m European funded Growth Catalyst programme, implemented by the University of Liverpool Management School is designed to develop entrepreneurial leadership competences as a stimulus for small firm growth in Liverpool. Informed by research to maximise entrepreneurial learning, the 8-month programme includes an initial 2-day Leadership Retreat, Masterclasses, Business Exchanges and individual Coaching. Between 2014 and 2018, 144 entrepreneurs from 129 small firms in 6 cohorts successfully completed Growth Catalyst.

The aim of this study is to evaluate the investment effects of this program on its initial cohorts for which sufficient data is available. Investment is a key proxy for the effectiveness of such programs for two reasons. First, successfully financing and implementing investment projects indicates improved management practices in the firm. Second, investment itself is a source of firm growth and productivity improvement. Programs generating large effects in terms of investment are likely to contribute to regional convergence.

The key issue when estimating causal effects of such programs is to handle self-selection resulting from more growth-oriented, ambitious and better-managed firms being more likely to participate. In the lack of experimental variation, the most credible econometric approaches design control group from firms which had performed similarly pre-treatment. For example, if treated firms had been growth-oriented even before the treatment, the control group should include similarly growth-oriented firms.

In addition, our approach attempts to answer two further issues. First, the small sample size typical in such trainings is not conducive to large-sample econometric techniques, such as panel regressions. Second, the training itself will benefit from firm-level, rather than only average, estimates because it allows trainers to reflect on what types of firms and managers are benefitted actually from the training.

Given these challenges, we opted to apply synthetic control methods to estimate the investment effects of the program (Abadie et al, 2010; Abadie et al, 2015). This method was designed to provide causal inference in comparative case studies. For example, in Abadie et al (2010) one US Stata, California, implemented a new policy on tobacco products, and the method was used to create a credible control group for California from the other 49 states. Clearly, the method is designed to cases when few units are treated.

The basic idea of the synthetic control group modelling is to create an artificial control group from the potential controls (called donors in this literature) by assigning weights to each of them. The weights are calculated in such a way that the synthetic control group (the weighted average of the donors) follows the treated unit as closely as possible pre-treatment. The causal effect is calculated as the difference between the treated unit and the synthetic control group after the treatment.

We rely on the FAME data on all UK firms in our analysis. Participating in the Growth Catalyst training is the treatment. For each of the firms we create a control group from firms in Merseyside and Cheshire (where the Liverpool city-Region is located) in the same 4-digit SIC industry. We calculate the weights in such a way that the synthetic control group follows the most similar trajectory in terms of fixed assets in the 4 years before the training and the year of the training. The treatment effect is estimated as the difference in growth rates of fixed assets between the treated firm and its control group during the 3 years after the training. This exercise allows us both to estimate the effect on each firm and the total effect of the program on this group of firms.

We build on the logic of Abadie et al (2010) to quantify the uncertainty of our estimates. In particular, we estimate a 'placebo' treatment effect for all firms in the sample, which allows us to calculate a distribution of the estimated effect on untreated firms. This can be used for exact inference.

We find that the Growth Catalyst initiative had a significant effect on firm level investment, equivalent to about 38% increase of the fixed assets for surviving firms on average. In monetary terms, this is equivalent to £283 thousand per firm or £36.5 million for the full Growth Catalyst program between 2014 and 2018.

In what follows, Section 2 describes the data we use, Section 3 our method. Section 4 summarizes the results and Section 5 concludes.

2 Data

Our main data source is the FAME company-level data of UK firms, which is the UK version of ORBIS/AMADEUS. It consists of key information for all UK firms, including their different identification numbers, address and industry. It also includes key financial information for a number of years. A limitation of the FAME data is that it usually includes only few financial variables for small firms. Importantly for our pruposes, fixed assets (and a few other measures of capital, including current assets) are available for most small firms. We will interpret the growth of fixed assets as a proxy for investment. Other variables, such as employment, sales, profits are available only for few small firms.

In terms of timing, data is available until 2018 for most firms. As a result, we can mainly analyse the medium-term effects of the cohorts which received their training in 2014/2015 where we have 3 post-treatment years (and to some extent those which participated in 2015/16 program, where we have 2 post-treatment years). We restrict our attention to firms for which data on fixed assets is available for all years between 2010 and 2018. We also restrict our sample to firms which operated either in Merseyside or Cheshire where Liverpool City-Region is located.

We link the firms which participated in the training to the FAME data with their registered company numbers. Altogether, 50 firms participted in the program in 2014/15. From these, we can link 49 to the FAME data. From these 16 were not active as of 2020.[1] From the remaning 33, 7 did not report its fixed assets every year between 2010 and 2018. This leaves us with a sample of 26 firms. In the 2016/17 cohort new 23 firms participated in the training. From these, 3 are not active anymore and 10 more did not report fixed assets for the full period. This leaves us with 10 observations from this cohort. Managers from 74 firms participated in the training in 2017/18.

3 Method

The first step of our method is to create the appropriate control group for each firm. We apply the methodology of Abadie et al (2010) modified for our purposes.

As a first step, we restrict the "donor set" to firms in the 4-digit SIC industry (which themselves did not participate in the program in any year) of the treated firm. From these we create the synthetic control group by minimizing the distance between the ln fixed assets of the treated firm and the synthetic control group in the (financial) year of the treatment and the 4 preceding years.[2] In other words, the control group is created in such a way that its investment patterns follow closely that of the participating firm for 5 years before and during the treatment. Technically, the control group is a convex combination of the firms in the donor pool, where $w_i \geq 0$ is the weigh of firm i and $\sum_i w_i = 1$.

---

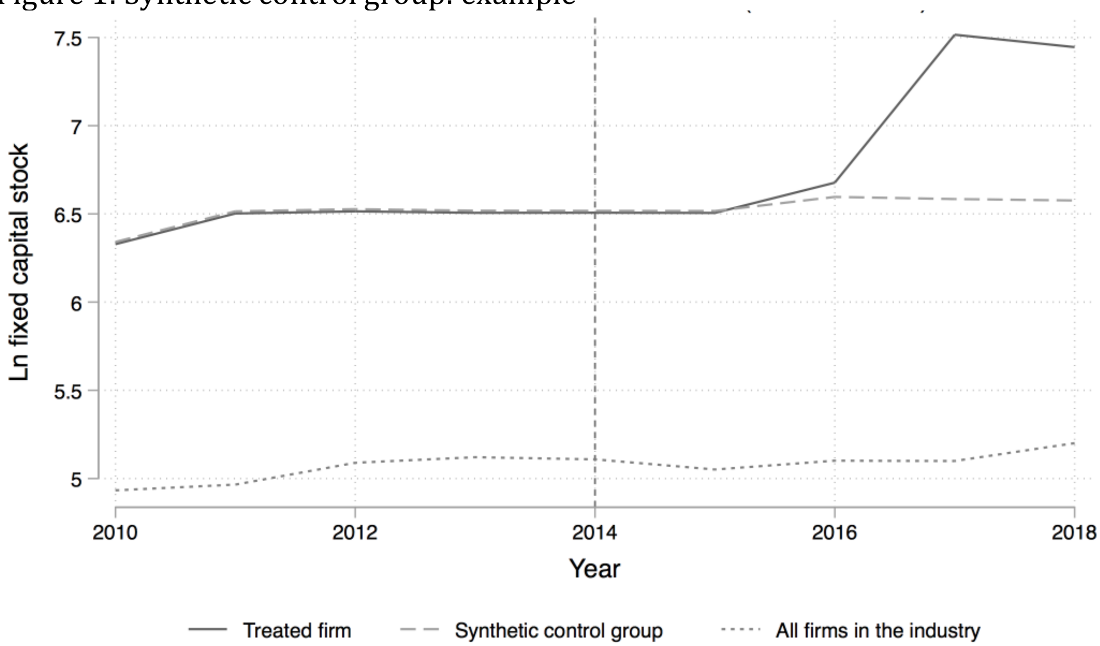
[1] The overall exit rate in the Northwest was 13.5 % per year in 2018 (ONS, 2019). This implies a 5-year survival rate (2015-2020) of 49% compared to the 67% survival rate in this sample. That said, our analysis omits the issue of differential survival rates and focuses only on continuously operating (and reporting) firms.
[2] Including the treatment year implicitly assumes no effect on investment in that year. We find this a conservative and plausible assumption. Matching only for years before the treatment yields substantially larger estimates. Effective synthetics control designs require relativelly long time periods before the treatment, which may be multiple decades when applied to countries and states. For small firms – and given our data limitations – the 5 year pre-trend seem to be a good compromise. We tested the robustness of the results by using only 3 years which did not change the results in important ways.

For each treated firm we can estimate the investment effect of the programme by comparing the growth of fixed assets in the treated firm to its growth in the synthetic control group relative to the year before the training. By denoting the time of the training with t_0 and the treated firm with j, we can estimate the investment effect of the program in T years by:

$$effect_j^T = \left[\ln fixedass_{j,t_0+T} - \ln fixedass_{j,t_0-1}\right] - \left[\sum_i w_i\left(\ln fixedass_{j,t_0+T} - \ln fixedass_{j,t_0-1}\right)\right]$$

Let us illustrate this with an example (Figure 1). The horizontal axis shows the different years. The training for this firm took place in the 2014 financial year. The trajectory of the synthetic control group follows that of the example firm closely until 2014. The fixed capital stock was still similar in 2015, showing that the program did not lead to immediate results, but the two series started to diverge from 2016, with substantial extra investment in the example firm in 2017. The extra investment of the example firm was 140% of its original fixed assets until 2018, which is shown by the difference between the two lines in 2018. As can be seen, the synthetic control group differs from the industry average somewhat both in terms of its capital level and investment.

Figure 1: Synthetic control group: example



Notes: This graph illustrates the working of the synthetic control group method on the example of one of the firms, treated in the 2014 financial year. The solid line shows stock of fixed capital of the treated firm in the different years. The dashed line shows the fixed capital stock of the synthetic control group. The two moves strongly together until 2014, when they start to diverge. The estimated effect is shown by the difference between the two lines in 2018. The dotted line is the industry average.

This method allows us to calculate point estimates for the effect of the training on each firm separately and also for the full program. For the full program, we simply add up the estimated effects:

$$totaleff^{T} = \sum_{j} effect_{j}^{T}$$

Naturally, the point estimates are estimated quantities and, therefore, they carry an uncertainty. By its very nature, the synthetic control method does not allow one to calculate straightforward confidence intervals like a regression model. Instead, following the logic of Abadie et al (2010) one can quantify the uncertainty of the estimate by conducting placebo tests on the donor pool. In particular, we repeat the whole procedure (generating a synthetic control group, calculating the treatment effect) for all of the non-treated firms and can calculate the "effect" of the non-treatment for all of these firms separately. If the distribution of these estimated placebo effects is similar to the estimated effect, then the treatment has no effect. If, on the contrary, these distributions differ substantially, the treatment is effective. Technically, the distribution of these estimated placebo effects provides an exact distribution of the treatment effect under the null of no effect.

We calculate these placebo effects for all firms which operate in the same industry as at least one of our treated firms and which had a tangible capital stock between the 5th and 95th percentile of the treated firms' distribution. This guarantees that these placebo firms are similar to our firms, mostly by excluding very large firms from our analysis.

We also use these placebo effects to quantify the uncertainty regarding our estimated total effect. We take 1,000 samples, each with the same number of observations as our treated firms, from the placebo effects and add them up. This yields an exact distribution for the total effect under the null of no effect.

This method is both intuitive and rigorous. However, given the non-experimental data set, it relies on some assumptions. Most importantly, it assumes "parallel trends", i.e. that the treated firm would have performed similarly to the control group without the treatment. This can be violated in a number of ways. One possibility is that firms decide to participate based on their future growth opportunities, which are not entirely captured by past growth patterns. Second, the firms may participate in other programs besides the Growth Catalyst, which may also affect their performance. Both of these possibilities may bias the estimated program effects upwards.

4 Results

4.1 Firm-level effects

Table 1 shows the average treatment effects for treated and non-treated (placebo) firms. We find no immediate effect of the program, while after three years the average effect of the program is 37.8% percent of the initial capital stock. The effect increase in time: it is estimated to be 16.2% after 1 year and 25.5 percent after 2 years. This increasing trend of the estimates is in line with the expectation that the program affects firm growth rather than generating a one-time jump.

Also in line with expectations, we don't find positive effects for the untreated firms, with slightly negative average placebo effects in each of the years. The last row shows p-values from 2-sided p-tests which compare the treated and placebo averages. We find that the differences are significant at the 10% level for the 1 and 2 year effects and at the 5% level for the 3-year effects. Given the small size of the treated sample, these provide strong evidence for the positive effects of the program.

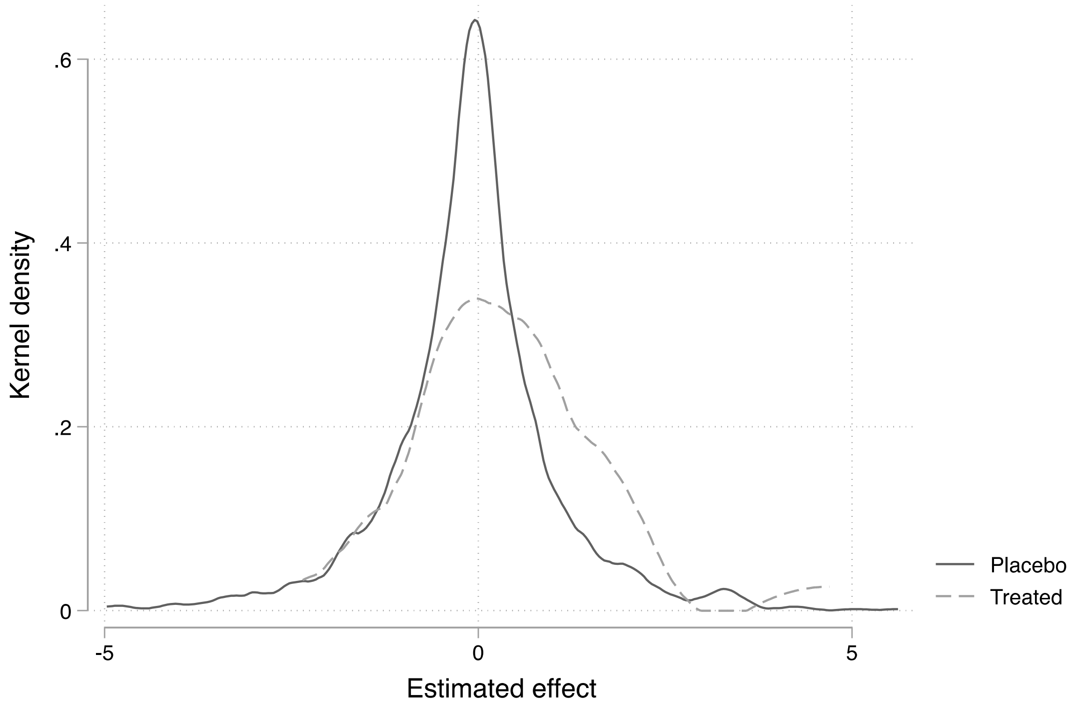
Table 1: Average estimated treatment effects (% of original capital stock)

Year:	immediate	1-year	2 years	3 years
Placebo	-2.4%	-6.4%	-8.9%	-10.2%
Treated	-2.3%	16.2%	25.5%	37.8%
P-value of t-test	0.49	0.08	0.06	0.02

Notes: This tables shows the mean estimated effects of the program (Treated) and the mean estimated placebo effects, estimated for non-treated firms with the same method. The average estimated effect after 3 years is 37.8%, which is different at the 5% significance level from the placebo estimates.

Besides the means, the whole distribution is also of interest. Figure 2 shows the distribution of the estimated 3-year effects both for the treated and non-treated (placebo) firms. The placebo distribution is centered slightly below zero. The treated distribution is asymmetric, with much heavier right tail. This shows that, while not all treated firms invested more than typical untreated firms, a number of treated firms invested substantially more.

Figure 2: Distribution of the 3-year estimated effects for the treated and non-treated (placebo) firms



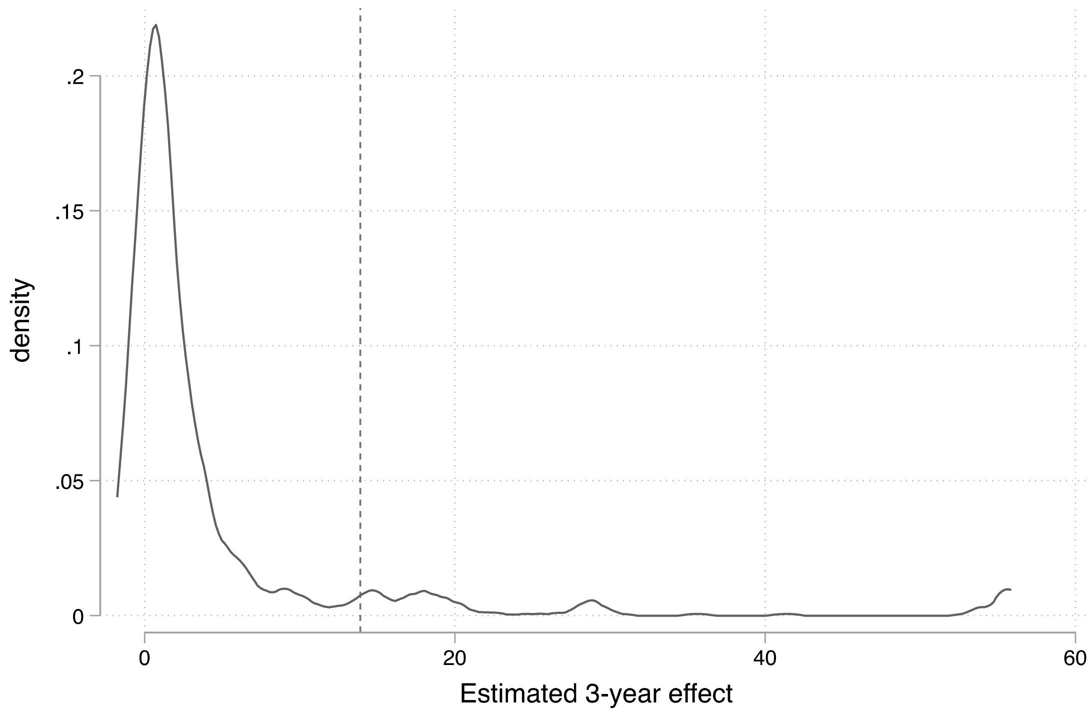
Note: This graph shows the kernel distributions of the estimated effects for treated firms and the placebo group. It shows substantial overlap, but there are clearly more substantially growing firms in the treated group.

4.2 Total effect

We can simply calculate the total effect of the training for the 26 firms in our sample by adding up their individual estimated effects, converted to £s. We find that the average extra investment was £536 thousand per firm, yielding a total estimate of £13.9 million for the 26 firms.

To quantify the uncertainty of this estimate we calculate the distributions of similar sums for the placebo firms. Figure 3 shows these estimates together with our estimate from the treated firms (dashed line). The estimate is clearly different from typical values from the placebo distribution, equal to its 93rd percentile.

Figure 3: Kernel distribution of placebo estimates for the total program together with the point estimate of £13.9 million (dashed line)



Note: This graph shows the estimated total effect of the program on the 26 firms in our main sample (dashed line, £13.9 m) and the distribution of total effects from the placebo firms. 93% percent of these are below the estimated total effects.

As we have discussed, altogether 49 firms were treated in the 2014/15 cohorts of the Growth Catalyst programme and only 26 was available in our data. As some of the other firms exited or entered in our sample period or did not supply positive capital data for some of the years, one may conservatively assume that the program did not have an effect on these firms. Under this assumption, the extra investment generated by the 2014/15 cohorts equal to the previous total of £13.9 million, with an average of £283 thousand per firm.

The full program trained managers from 129 firms between 2014 and 2018. Assuming that the average, £283 thousand, is valid for all cohorts of the program, we can estimate that the full effect on these cohorts may be around £36.5 million.

5 Conclusions

This paper have adapted the syntheitc control group method for evaluting the impact of the Growth Catalyst program. The method seems to be well fitted to evaluating and providing feedback for such programs because it works well when there are few treated units and provides estimates for each firm.

We find that the Growth Catalyst initiative had a significant effect on firm level investment, equivalent to about 38% increase of the fixed assets for surviving firms on average. In monetary terms, this is equivalent to £283 thousand per firm or £36.5 million for the full Growth Catalyst program between 2014 and 2018.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American statistical Association*, *105*(490), 493-505.

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495-510.

Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics*, *122*(4), 1351-1408.

ONS (2019). **Business demography, UK: 2018,** Change in the number of UK businesses broken down by sector of the economy. Available at: https://www.ons.gov.uk/businessindustryandtrade/business/activitysizeandlocation/bulletins/businessdemography/2018