Employing the unemployed of Marienthal: Evaluation of a guaranteed job program

PRE-ANALYSIS PLAN

Maximilian Kasy Lukas Lehner

November 4, 2020

This documents outlines the study design and statistical analysis of our evaluation of a guaranteed job program that will be piloted in the municipality of Gramatneusiedl in Austria, starting in October 2020. We first provide some background on the program and intervention, including the timeline of the job guarantee pilot. We then discuss our study design, which is based on two approaches. The first approach uses individual-level pairwise randomization of participants into waves. The second approach constructs a synthetic control at the municipality level, to impute counterfactual outcomes for Gramatneusiedl. We consider multiple different treatment contrasts to separate out direct effects of program participation, anticipation effects of future participation, and municipal level equilibrium effects. This pre-analysis plan concludes with a discussion of our outcomes of interest, hypothesis testing based on randomization / permutation inference, as well as multiple-testing corrections.

1 Background

1.1 Description of the intervention

Starting in October 2020, the Public Employment Service for Lower Austria (*Arbeitsmarktservice Niederösterreich (AMS NÖ)*) is piloting an intervention that aims to eradicate long-term unemployment and improve social, health and wellbeing outcomes for people in long-term unemployment, by bringing them back into employment. The intervention will provide a guaranteed job complemented by targeted counseling to support people in long-term unemployment.

The intervention will take place in one town in Lower Austria, Gramatneusiedl. All residents who have been unemployed for over 9 months are eligible to participate. The initial period for the project is set until 2024 and budgeted with EUR 7.4 Million for the full duration. The *AMS* calculates that the annual cost of the intervention is EUR 29,841.39 per participant.

Preparatory training The program is implemented by the private service-provider *itworks*, which specializes in implementing active labour market programmes for the *AMS*. *itworks* provides preparatory training for participants, and continues counseling and training after participants have taken up employment. The preparatory training phase is scheduled for a maximum of eight weeks, but durations may vary depending on individual conditions and progress. Each participant receives a tailored curriculum according to their individual needs. This may include individual and group counseling, skills development, supporting self-initiative, and assistance with health-related problems. Participants will continue to be encouraged to take up regular employment outside of the program if available.

Guaranteed jobs After successful completion of the preparatory training phase, participants join the job guarantee program for up to 3 years. Participants are supported to find a job on the regular labor market, for which the *AMS* subsidizes wage costs at a 100% rate for the first 3 months, and at 2/3 for the subsequent 9 months. Those that remain without job placement will receive an employment offer with a newly established social enterprise operated by *itworks*. All participants will be paid at least minimum wage. The social enterprise implements projects at the municipal and regional level. Tasks may include activities such as working in childcare, gardening, renovation, and carpentry depending on orders acquired by the enterprise. In addition, participants are supported to develop and propose their own ideas for projects of the social enterprise based on their expertise and local knowledge of community needs. Through its business activities, the enterprise is expected to generate revenues of around EUR 383,000 over the project duration.

A specific effort is made to create productive and meaningful employment that is adequate to the participants' previous jobs in terms of skills and income. The jobs created are tailored to the needs of the recipients. Persons with limited work availability receive targeted job offers. A person only available to work part-time receives a corresponding part-time offer. A person who can carry out only a limited number of tasks (e.g. for health reasons) similarly receives a corresponding offer. Social workers and instructors continue to provide support to employees of the social enterprise as needed. Participants have access to occupational physicians. Those participants that feel ready to work for third-party employers receive targeted support and additional counseling to apply and find employment outside of the program.

Voluntary participation Work conditionality is eased. Currently, under law (*Arbeitslosenversicherungsgesetz (AlVG §9)*), people in unemployment are assigned to labour market programs by the *AMS*. They have the obligation to participate and they have to accept an employment offer that conforms to their skillset. By contrast, within the job guarantee program only participation at the information event and the preparatory training phase is subject to this conditionality, while take-up of employment offered as part of the job guarantee is voluntary (i.e.. without sanctions in case a job offer is declined).

1.2 Timeline for the intervention

- October 2020: Targeted curriculum and coaching starts; (first wave of 31 participants).
- December 2020: Intended start of employment (first wave of 31 participants).
- February 2021: Targeted curriculum and coaching starts; (second wave of 31 participants).
- April 2021 Intended start of employment (second wave of 31 participants). The treatment continues for (at least) 3 years up to March 2024.

A complementary study conducted by researchers at the Department of Sociology, University of Vienna will use a mixed-methods design and qualitative in-depth interviews, including longer-term follow ups. Data collection for the respective studies will be timed with several months in between, so as not to overburden participants.

1.3 Impact of the COVID-19 pandemic

Due to the COVID-19 pandemic, labor market conditions have worsened in Lower Austria, including Gramatneusiedl. The trajectory of Gramatneusiedl was similar to comparable municipalities as indicated in the synthetic control comparison. All individuals included in our treatment and control groups for the pairwise matching have become unemployed before the pandemic. However, their opportunities to find employment since have been severely impacted. Entrants into the job

guarantee scheme at a later stage - relevant for the synthetic control comparison - will include those who became unemployed during the pandemic.

The implementation and timeline of the job guarantee pilot were not affected, and the pilot continues as planned. We plan to take specific precautionary measures during the fieldwork and data collection to guarantee the safety of both the participants and the researchers involved. We have detailed those in the ethics application for our study that was approved by the Departmental Research Ethics Committee at the Department of Economics, University of Oxford.

2 Study design

2.1 Overview

Our evaluation of the job guarantee program is based on two complementary study designs. The first design uses pairwise randomization within pairs of participants matched using baseline covariates; cf. Athey and Imbens (2017). Participants are assigned to one of two waves, where the second wave starts the program 4 months after the first one. This allows us to estimate (short-term) effects of the program by comparing participants across the two waves, around 3-4 months from the start of the first wave.

The second design uses the synthetic control method; cf. Abadie et al. (2010). We construct a synthetic control town for Gramatneusiedl, based on other towns in the province of Lower Austria. The synthetic control town is a convex combination of similar towns. This method allows us to estimate effects of the program at the town level, including potential spillovers on non-eligible residents.

The code implementing the following designs has been uploaded to GitHub, at `https://github.com/maxkasy/Marienthal`. For the matched pair design, we use the package *nbpMatching* in R, for the synthetic control design we use the package *Synth*.

2.2 Sample selection

The set of participants who are eligible for the job guarantee program includes all current residents of Gramatneusiedl registered with the *AMS* who are at risk of long-term unemployment or in long-term unemployment. "At risk of long-term" means an unemployment spell exceeding 9 months and "long-term" means an unemployment spell exceeding 12 months. The definition of unemployment used here is "beschäftigungslos". This implies that the duration of unemployment is counted irrespective of whether individuals have participated in active labor market programs of the *AMS* during their unemployment spell. It also includes those who have registered sick leave for less than 62 consecutive days or have attempted to take up employment but were employed for less than 62 consecutive days within the last 9 months. Only if a formerly unemployed person returns to unemployment from sick leave or employment that lasted longer than 62 days, the count of the unemployment spell duration starts again from zero.

The set of municipalities used to construct our synthetic control municipality includes all municipalities in the province of Lower Austria.

2.3 Assignment of participants to waves

As noted above, we assign program participants to one of two waves using pairwise randomization, where pairs are matched using a number of covariates. These covariates include the following, which were constructed from the raw data for the eligible participants provided by the *AMS*: Gender, age, "migration background" (being a migrant or child of migrants), education (more than "Pflichtschule," the legally required minimum), presence of a disability or health condition

recorded by the *AMS*, the level of benefits most recently received (closely correlated with prior income), and the number of days recorded as unemployed and looking for a job in the last 10 years. All of these variables are used as available to the *AMS* at the latest available date in September 2020. These data have been recorded at the last prior interaction between each of the participants and the *AMS*.

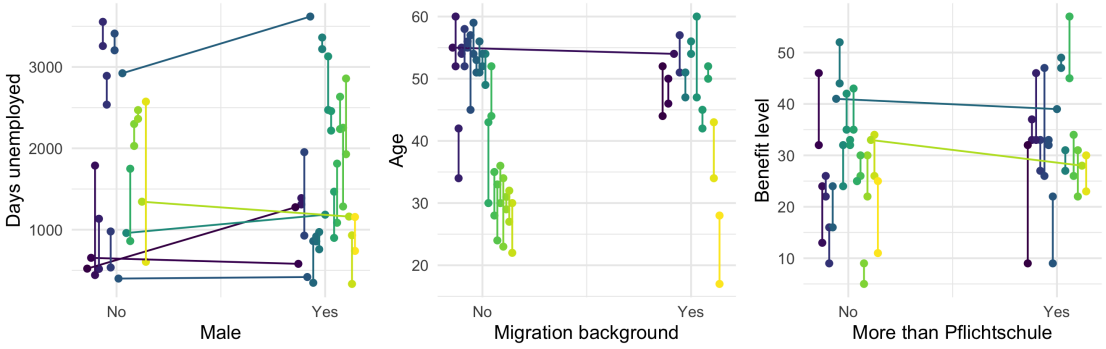
Pairwise distances between all 62 program participants are calculated using the Mahalanobis distance[1] based on these covariates. Participants are matched into pairs such that the total sum of distances between the members of each matched pair is minimized. One of the participants in each pair is then randomly selected to start the program in wave 1, while the other participant is assigned to wave 2.

Summarizing the resulting assignment, Table 1 shows the differences in covariate means between waves, and the corresponding t-statistics. Confirming that our procedure worked as intended, all available covariates are balanced across waves. Providing a more disaggregated representation of the matching, Figure 1 depicts the matched pairs in terms of the different covariates characterizing participants.

Table 1: Covariate balance for our matched pair design

Covariate	Mean wave 1	Mean wave 2	Difference	T-statistic	P-value
Male	0.581	0.581	0.000	0.000	1.000
Age	44.452	44.935	-0.484	-0.165	0.869
Migration Background	0.323	0.355	-0.032	-0.264	0.793
Education	0.452	0.452	0.000	0.000	1.000
Health condition	0.290	0.323	-0.032	-0.271	0.787
Benefit level	29.839	29.839	0.000	0.000	1.000
Days unemployed	1721.871	1600.839	121.032	0.483	0.631

Figure 1: Matched pairs



[1]The Mahalanobis distance of two covariate vectors x_1 and x_2 that are realizations of a random vector X is given by $d(x_1, x_2) = \sqrt{(x_1 - x_2) \cdot Var(X)^{-1} \cdot (x_1 - x_2)}$.

2.4 Construction of the synthetic control

We next turn to our second approach, constructing a synthetic control municipality for Gramatneusiedl. For the construction of this control municipality we draw on data from various sources, including the *AMS* internal registry (*"AMS Datenbank"*), the social security registry (*AMS BMAFJ Erwerbskarrierenmonitoring* or *AMDB*), and the national statistical agency (*STATcube - Statistische Datenbank* of *Statistik Austria*). All data were retrieved in September 2020.

We construct a synthetic control municipality in two steps. In the first step, we select a subsample of 5% of the available municipalities (26 out of 505 municipalities) which are most similar (closest) to Gramatneusiedl. Similarity is measured in terms of the Mahalanobis distance in covariate space. The covariates used are listed in Table 3 below. All covariates are based on observations for the year 2019 (as measured in December). Additionally, we also include some indicators measured in July of 2020, after the onset of the Covid pandemic, to account for possibly heterogeneous impacts of this pandemic. These covariates are shown in the last panel of Table 3.

In the second step, we construct a synthetic control based on these 26 municipalities, using the approach described in Abadie et al. (2010) and reviewed in Abadie (2019). This synthetic control is chosen to match the same list of covariates used in the first step (where we selected a subsample of municipalities), as well as additionally the trajectory of unemployment rates (unemployed as a share of the working age population) in Gramatneusiedl from 2011 to 2020, that is, for the 10 years preceding the intervention.

The resulting weights are shown in Table 2, which lists all municipalities with non-negligible weights. The time series of the predicted unemployment rate using the synthetic control, and the corresponding realized time series of unemployment for Gramatneusiedl in the 10 years preceding the intervention are shown in Figure 2. Table 3 compares the covariate values for Gramatneusiedl with those for the synthetic control as well as those for each of the municipalities with positive synthetic control weights.

Variable definitions The variables used to select similar municipalities and to construct the synthetic control, as shown in Table 3, are defined as follows:

- unemployed as a share of working age population,
- long term unemployed as a share of working age population,
- working age population,
- inactive population as a share of working age population,
- mean age of the total population,
- small firms (less than 10 employees) as a share of total firms,
- medium sized firms (10-249 employees) as a share of total firms,
- population with low education (ISCED 1-2) as a share of total population,
- population with medium education (ISCED 3-4) as a share of total population,
- male population as a share of total population,
- population with a migrant background as a share of total population,
- active population with care responsibilities as a share of total population,
- mean wage level,

- mean age of the unemployed,
- unemployed with low education (ISCED 1-2 as a share of total unemployed,
- unemployed with medium education (ISCED 3-4) as a share of total unemployed,
- unemployed with low German skills (less than A2 CEFR) as a share of total unemployed,
- male unemployed as a share of total unemployed,
- unemployed with a migrant background as a share of total unemployed,
- unemployed with an underlying health condition limiting employment opportunities as a share of total unemployed,
- communal tax per working age population.

Table 2: Synthetic control weights

Weight	Municipality	Identifier
0.487	Ebreichsdorf	30607
0.203	Zeillern	30544
0.134	Rußbach	31224
0.079	Leopoldsdorf im Marchfelde	30831
0.046	Strasshof an der Nordbahn	30856
0.024	Sieghartskirchen	32131
0.023	Sollenau	32327

Figure 2: Synthetic control gap

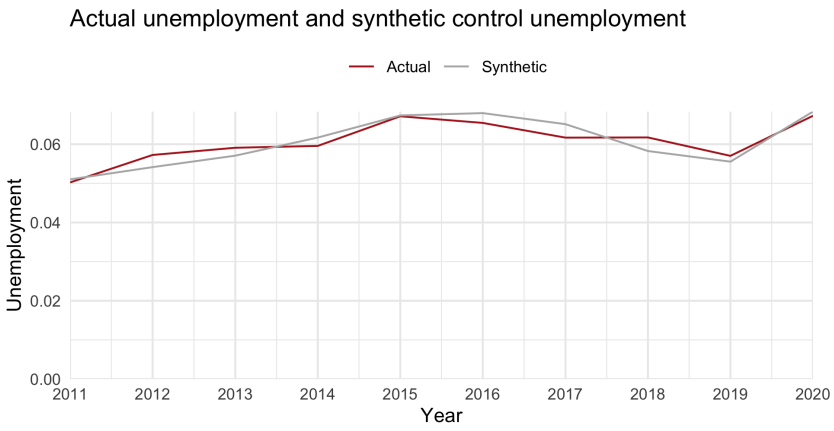


Table 3: Gramatneusiedl and control municipality covariates

Municipality	Working age pop	Long term unemp/pop	Inactive/pop	Mean age	Share small firms	Share mid firms	Share low edu
Gramatneusiedl	5013	0.007	0.220	50.775	0.115	0.339	0.208
Synthetic control	4830	0.016	0.228	51.074	0.126	0.363	0.225
Zeillern	1263	0.004	0.227	50.229	0.093	0.335	0.199
Ebreichsdorf	7655	0.020	0.228	50.810	0.139	0.381	0.235
Leopoldsdorf im Marchfelde	2035	0.022	0.247	51.304	0.135	0.348	0.242
Strasshof an der Nordbahn	6920	0.024	0.213	51.403	0.115	0.324	0.250
Rußbach	942	0.013	0.219	52.230	0.126	0.369	0.206
Sieghartskirchen	4560	0.010	0.224	52.464	0.135	0.337	0.197
Sollenau	5122	0.017	0.248	54.286	0.129	0.360	0.284

Municipality	Share mid edu	Share men	Share migrant	Share care resp	Mean wage	Mean age unemp	Low edu/unemp
Gramatneusiedl	0.642	0.511	0.242	0.257	3416	42.694	0.530
Synthetic control	0.644	0.503	0.181	0.235	3293	43.422	0.452
Zeillern	0.702	0.509	0.053	0.256	3168	40.462	0.346
Ebreichsdorf	0.620	0.498	0.234	0.235	3379	44.344	0.465
Leopoldsdorf im Marchfelde	0.619	0.498	0.260	0.216	3294	43.627	0.513
Strasshof an der Nordbahn	0.600	0.496	0.276	0.257	3393	42.364	0.465
Rußbach	0.676	0.513	0.088	0.224	3137	45.500	0.525
Sieghartskirchen	0.641	0.510	0.195	0.206	3366	41.257	0.387
Sollenau	0.608	0.496	0.229	0.193	3235	41.819	0.521

Municipality	Mid edu/unemp	Poor German/unemp	Men/unemp	Migrant/unemp	Health cond/unemp	Communal tax/pop	Lt ue/pop 2020
Gramatneusiedl	0.455	0.082	0.627	0.418	0.245	57.281	0.009
Synthetic control	0.516	0.061	0.583	0.312	0.264	217.301	0.018
Zeillern	0.654	0.000	0.692	0.115	0.303	97.822	0.004
Ebreichsdorf	0.480	0.086	0.546	0.374	0.213	282.242	0.022
Leopoldsdorf im Marchfelde	0.473	0.093	0.573	0.467	0.256	284.806	0.023
Strasshof an der Nordbahn	0.496	0.089	0.528	0.472	0.303	160.549	0.027
Rußbach	0.475	0.025	0.575	0.200	0.375	97.079	0.016
Sieghartskirchen	0.552	0.054	0.609	0.360	0.281	329.855	0.012
Sollenau	0.460	0.140	0.558	0.457	0.282	308.998	0.019

2020

Municipality	Inactive/pop	Mean wage	Mean age ue	Low edu/ue	Mid edu/ue	Poor German/ue	Health cond/ue
Gramatneusiedl	0.209	3308	42.069	0.456	0.481	0.031	0.209
Synthetic control	0.219	3181	42.625	0.389	0.577	0.059	0.212
Zeillern	0.222	3025	41.474	0.289	0.711	0.000	0.193
Ebreichsdorf	0.217	3278	43.101	0.424	0.527	0.082	0.169
Leopoldsdorf im Marchfelde	0.244	3222	44.021	0.472	0.507	0.056	0.225
Strasshof an der Nordbahn	0.202	3264	41.188	0.458	0.493	0.061	0.260
Rußbach	0.208	3022	42.314	0.343	0.629	0.057	0.349
Sieghartskirchen	0.220	3241	43.406	0.319	0.626	0.043	0.278
Sollenau	0.238	3071	41.847	0.460	0.517	0.119	0.274

3 Estimation and inference

3.1 Outcomes of interest

Contrasts In order to assess the impact of the guaranteed job program, we will consider three contrasts. First, we compare participants in the two waves of the pairwise randomized study. This comparison delivers credibly identified treatment effects. It is restricted, however, to short term outcomes measured in early 2021, before the second wave of participants starts their jobs. Furthermore, the control group might be impacted by the anticipation of future program receipt.

Second, we construct an additional control group from the long term unemployed who live in synthetic control municipalities, and who would have been eligible to participate had they been residents of Gramatneusiedl. This comparison relies on the validity of the synthetic control. It allows us, however, to estimate treatment effects which are not contaminated by anticipated program receipt and to compare outcomes over a longer period.

Third and lastly, we estimate municipal level treatment effects by comparing Gramatneusiedl to the synthetic control. This comparison includes residents who would not have been eligible to participate in the program, because they were not long-term unemployed. This comparison again relies on the validity of the synthetic control. It allows us to estimate equilibrium effects and spillovers at the municipality level, which might for instance be driven by the crowd-out of jobs, or by consumer demand effects of those participating in the program.

Primary outcomes We will estimate the effect of program participation on a range of economic and social outcomes. To assess the main objective of the intervention - to eradicate long-term unemployment - we will use labor market status as our primary outcome; for participants in the experimental study, and for other residents of Gramatneusiedl in the synthetic control study. For the experimental study, labor market status will be defined as being in employment vs. not being in employment. For the synthetic control study, we will estimate the treatment effect on the unemployment rate (unemployed as a share of the working age population), which we expect to exhibit a larger effect. For robustness, we will also estimate the effect on the employment rate.

Secondary outcomes To assess the impact of treatment on secondary objectives, we will estimate the treatment effect on a range of outcomes, which are provided in Table 4. We will use multi-item scales to construct one composite index for each of the secondary outcomes, except for employment and income. We then group those outcomes in four thematic families. We will register the full questionnaire and methodology for constructing the composite indices before running the first survey.

Administrative and survey data Data on labor market status and income will be available from the same administrative data sources that we used to design the study, as outlined in Section 2. We will collect all other data by surveying the experimental participants in both the treatment and control groups, as well as the program instructors. Additionally, we will survey individuals in control towns, who would meet the elibility criteria for program participation (i.e., are unemploymed for longer than 9 months).

Attrition We will make an effort to keep attrition to a minimum, ensuring that all those eligible to participate are included in our surveys. Nevertheless, we expect some of the individuals to attrit, either because they refuse to participate in the data collection for our study, or because they move abroad, or pass away. We will report intention to treat effects and bounds accounting for potentially selective attrition.

Table 4: Secondary outcomes

Economic outcomes	Health	Social outcomes	Social capital
labor market status	physical	material deprivation	trust[4]
income[1]	psychological	wellbeing	civic engagement
skills[2]		social inclusion	attitudes[5]
		general condition[3]	

[1] includes unemployment benefits and wages.
[2] employment related.
[3] includes appearance, hygiene, and time management.
[4] in society and institutions.
[5] towards the welfare state, migration, and environmental protection.

3.2 Inference for the pairwise randomized sample

Randomization inference Our approach toward inference is based on permutations of treatments, that is, based on randomization inference. This approach allows us to test the null hypothesis that the intervention had no effect of any kind, that is, $Y_i^1 = Y_i^0$ for all individuals i and potential outcomes Y_i^1, Y_i^0.

We re-assign treatment at random *within* (!) each of the matched pairs of participants. For this counterfactual treatment assignment, we re-calculate any given test-statistic. Repeating this process many times, we can calculate the share of re-assignments for which the test-statistic is bigger than the realized value of the test-statistic. This share is the p-value for the null hypothesis of no effects.

Compound hypotheses In order to deal with the issue of multiple testing in a principled manner, we will use the Benjamini–Hochberg procedure, which allows us to control the false discovery rate, that is, the share of rejected hypotheses which in fact hold true.

This procedure works as follows. Sort the p-values, for each of the m hypotheses, tested by size, resulting in ordered values $P_{(j)}$. For a critical value α, find the largest value k such that

$$P_{(k)} \leq \frac{k}{m}\alpha.$$

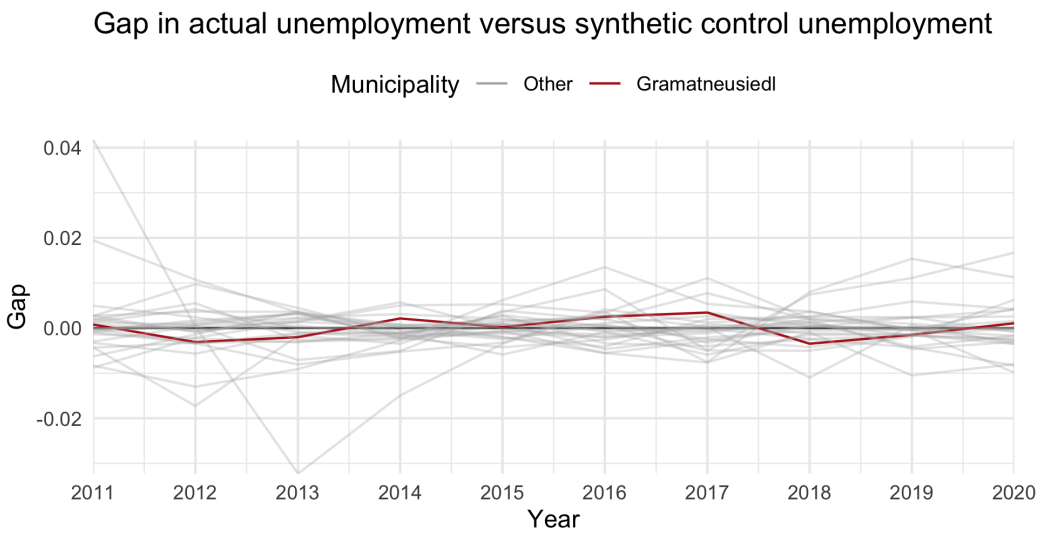
Reject the null hypothesis for all $i = 1, \ldots, k$.

3.3 Permutation inference for the synthetic control

Our inference for the synthetic control method relies on the permutation approach as described in Abadie et al. (2010). We consider Gramatneusiedl and each of the 25 control municipalities based on which the synthetic control for Gramatneusiedl was constructed. For each of these, we calculate a synthetic control based on the other 25 municipalities, and use this synthetic control to predict outcomes in the post-intervention period. The share of these municipalities for which the resulting gap between realized and predicted outcomes is larger than for Gramatneusiedl can then be interpreted as a p-value for the null-hypothesis that the intervention had no effect on the outcome for Gramatneusiedl.

Figure 3 shows the time series of gaps between realized and predicted outcomes for Gramatneusiedl and for the control municipalities during the 10 years before the intervention. As can be seen from this figure, Gramatneusiedl is not an outlier during this pre-period.

Figure 3: Permutation inference for the synthetic control



4 Publication agreement

This evaluation is based on an agreement between the researchers (us, Lukas Lehner and Maximilian Kasy) and the *AMS NÖ*. Two key components of this agreement are that (1) no payment will be made from the *AMS NÖ* to the researchers, and (2) the researchers are guaranteed to be entitled to publish the findings of their study in academic outlets without any interference by the *AMS NÖ*. The *AMS NÖ* in turn retains the right to disseminate the findings to the wider public, including to news outlets and social media.

5 IRB approval

The experimental work described in this pre-analysis plan was reviewed and approved by the Departmental Research Ethics Committee at the Department of Economics, Oxford University.

References

Abadie, A. (2019). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*.

Abadie, A., Diamond, A., and 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.

Athey, S. and Imbens, G. W. (2017). The econometrics of randomized experiments. In *Handbook of Economic Field Experiments*, volume 1, pages 73–140. Elsevier.