Oma linja school intervention: Preliminary pre-analysis plan

Tuomas Pekkarinen, Hanna Pesola and Matti Sarvimäki

December 22, 2020

This is a pre-analysis plan for the evaluation of the *Oma linja* -intervention conducted for 9th graders in Finland in 2016–2020. We document the intervention, randomization approach, power calculations, and a plan for register-based outcome and control variables and econometric methods.

1 Intervention

The *Oma linja* -intervention aims to help students at their final year of compulsory education (9th graders) to choose secondary education suitable for them. We examine two versions of the program. The "intensive" intervention consists of (a) motivational workshops and (b) implementation of social-cognitive group counselling techniques in the classrooms. The "light" intervention includes only group counselling activities.

1.1 Workshops

In the motivational workshops students are guided to identify their personal strengths and career aspirations. The presenters tell stories from their own lives with the aim of promiting self-exploration among students. The presenters participate in group works and help guide the discussions. The development of the group counseling techniques is based on earlier experience in group methods aiming at career choice preparedness (Vuori et al. 2008).

1.2 Group counselling

Student counsellors from each intervention school participate in two-day teacher training organized by the Finnish Institute of Occupational Health. During the training counselors are familiarized with the content and counselling principles of the intervention. After the training workshop, the counselling techniques are implemented at school. Implementation support will be provided throughout the year to schools. In addition, school counselors will receive structured guidelines and learning material to help carrying out the intervention. Group counselling techniques comprise four components:

1. *Career management skills training*. The students are supported to identify personal strengths and career interests, explore various occupational choices and set meaningful career goals and plans.

2. *Active teaching and learning methods*. Instead of lecturing, the trainers use the knowledge and experiences of the participants themselves as part of the learning process. Teachers activate and facilitate the learning process and guide the participants towards the desired conclusions.

3. *Supportive learning environment*. Principles of social learning (Bandura 1986) guide the student-centered approach, which includes mastery experiences through problem-solving exercises, learning vicariously and receiving peer reinforcement during group discussions.

4. *Inoculation against setbacks* (Meichenbaum 2017). Students are encouraged to analyze obstacles and setbacks they may face in their educational and occupational careers. The students are guided in problem-solving processes where they learn to cope with the stress related to education and career transitions. The purpose of inoculation against setbacks is self-preparation for coping with problems in career and school transitions due to lack of social support and guidance, conflicts with friends or parents and lack of confidence in one's own success.

2 Evaluation questions

Our aim is to provide a comprehensive analysis of the impacts of the *Oma linja* -intervention over the lifecycle of the participants. Our first set of results will examine effects on educational outcomes (application patterns, enrollment, drop-out, program changes, completion and grades in post-mandatory education). In a later stage, we plan to extend the analysis on the long-term impact on labor market, family, crime and health outcomes. This part of the analysis is contingent for data availability and research funding. Thus, we will file a separate pre-analysis plan in case we are able to carry out this analysis.

3 Experimental design

3.1 Target schools

Our target population consists of Finnish-speaking middle schools with high dropout rates located in 16 municipalities (Espoo, Helsinki, Hyvinkää, Hämeenlinna, Jyväskylä, Järvenpää, Kirkkonummi, Kotka, Lahti, Lappeenranta, Oulu, Porvoo, Salo, Tampere, Turku, Vantaa). We exclude speciality schools such as Steiner schools, schools exclusive targeted to disabled students and schools that do not have at least two parallel classes for grades 7–9. In total, our target population consisted of 92 schools that had 480 regular classes and approximately 8,400 ninth grade students.

3.2 Randomization

3.2.1 School-level randomization

We used randomized block design to divide schools into treatment and control groups. We created stratas using school-level dropout rates constructed with data from Statistics Finland's *Sijoittumispalvelu*. "Dropouts" were defined as individuals who had not obtained a secondary degree and were not enrolled in any school four years after graduation. Within each municipality, we ranked the schools according to their earlier dropout rates and divided them into bins of two schools. The randomization for the first two rounds were conducted simultaneously. For these rounds, we also randomized the bins into those treated during the 2016–2017 and 2017–2018 academic years. The motivation for this step was to ensure that we do not create correlation between the year the school was treated and school "quality". From each bin, we then randomized one school to become a treatment school, while the other one would become a control school. The school-level randomization of the remaining two rounds were conducted during the spring term of the academic year preceding the intervention.

3.2.2 Within-school randomization

Once the schools were recruited to participate, we randomized classes in the treatment schools into intensive and light treatment groups (see above). In the control schools, we randomized classes into a group for which we conducted the same survey as in the treatment schools and into "pure control", which was not contacted at all.

3.3 Power calculations

The sample size will allow us to detect effects of 2.4 percentage points change in graduation from upper secondary education four years after graduating from high school for power of 80% with a significance level of 5%. The power calculations are conducted by running simulations using data on students graduating from high schools in the treatment area in 2003–2008. We interpret the results of these power calculations to be conservative due to the small number of control variables in our simulation data and the *ad hoc* choice of specification. As we discuss below, we will have a much richer set of control variables and expect to take advantage of more sophisticated empirical methods at the time of conducting our final analysis. Thus we expect to have at least as much statistical power as implied by these rough power calculations.

4 Data

4.1 Primary data sources

Once the students leave 9th grade, we continue following them primarily using register data. These data will be constructed by Statistics Finland by linking together several administrative registers. These registers contain substantial amount of information on the population residing in Finland including detailed information on application, enrollment and graduation from post-mandatory education. Register-based data will be also available for labor market outcomes, crime, health, family formation and so forth. Furthermore, we will use register-base data to form control variables containing information about family background of the students participating in the intervention.

The availability of register-based data is constantly evolving and thus we are not able to provide a full description of the data that will be available at the time of evaluating the intervention. Statistics Finland's website includes descriptions of the currently available data.[1]

4.2 Outcome variables

As noted above, there remains some uncertainty concerning the precise content of register data that will be available at the time of evaluating the impact of the intervention. Thus we will discuss register-based outcomes as precisely as we can, but leave some ambiguity to the exact definitions of these outcome variables. We will update the pre-analysis plan if better data sources become available before we get access to data containing information on treatment and control groups.

- Our **primary outcomes** are:
  - **Graduation from upper-secondary education**. We use data from the Registry of educational degrees to examine first the effect of the intervention on the probability of graduating from secondary education three years after the intervention. We will then subsequently examine the effect of the intervention on graduating from secondary school four, five, six and seven years since leaving mandatory education as we gather more data. We measure this outcome by using a binary variable for having graduated from any upper-secondary program. In addition, we will create a measure of the quality of the secondary education using data on earlier graduates.
  - **Enrollment in tertiary education**. Similarly as above we examine the effect of the intervention on enrolling in tertiary education four years after the intervention and subsequently five, six and seven years later. Enrollment will be measured with a binary variable that takes value one if the individual is enrolled in any kind of tertiary education. As above, we will also create a measure of the quality of tertiary education using data on earlier graduates.
- Our **secondary outcomes** are:
  - **Application behavior of the subjects**. We use data from the applications registers to examine the effect of the intervention on the application portfolio of the subjects. Since there are several ways of characterizing application portfolios,

[1] At the time of writing, the information is available at www.stat.fi/tup/mikroaineistot/index_en.html.

we limit our choice of outcome variables by following (Goux et al. 2017) and examine the effect of the intervention on:

1. the probability of applying to post-mandatory education at all
2. the probability that the application portfolio contains at least one vocational programme
3. the probability that a vocational programme is ranked first
4. the probability that a vocational programme is not ranked first
5. the portfolio only contains academic programmes.

Furthermore, we will chracterize the secondary school programmes by calculating the predicted probabilities of graduating from each programme for all the individuals given their compulsory school grades and demographic characteristics. These probabilities are derived from regressions where graduation in each programme is regressed on school grades and demographics using pre-intervention data.

– **Enrollment in upper-secondary education during the two years after graduation**. We will use a binary variable for having been enrolled in any school during two years following graduation (e.g. for those graduating in spring 2017, the follow-up period will be from the fall term of 2017 to the fall term of 2019). We will also use a measure of the type of program the students are enrolled in. The likely alternatives in this regard include splitting the programs into preparatory studies, vocational track and academic track and/or using finer summary measures of the program type such as average outcomes of previous students enrolled in the school and/or program.

– **Enrollment in any education during the two years after graduation** This includes upper-secondary education as well as the supplementary classes of the comprehensive school.

– **Employment and earnings**. We obtain data on annual earnings and months of employment for all the years following the intervention. — We create an indicator for being neither employed or enrolled (NEET). We will examine the effect of the intervention NEET.

– **Program switches and grades in upper secondary education.**

- **Social transfers** Use of social transfers that are identifiable in the register data during ages 18-25.
- **Criminal activity** We use data on decisions by the district courts to examine the effect of the intervention on the propensity to commit a crime that leads to a conviction in a district court during ages 18-25. In addition, we will create a measure of the seriousness of the offences using data on previous cohorts. For this measure we will combine data from the district courts with data on offences and coercive measures registered by the police. We will specify this measure in more detail in a later update to this PAP.

5 Econometric approach

As noted above, there remains some uncertainty concerning future data availability at the time of writing both in terms of potential outcome and background variables. Furthermore, machine learning approaches for program evaluation are evolving rapidly and we hope to take full advantage of these opportunities in terms of specification choice, subgroup analysis and so forth. Thus we present only an initial discusion of empirical methods and outcome variables and may update the pre-analysis plan at a later stage.

5.1 Main specifications

Our main analysis will be at the assignment level and thus we focus on "intention-to-treat" estimates. Subject to data availability, we will discuss the extent to which the intervention reached the students assigned to each group and present local average treatment effect estimates using assignment status as an instrument for participating in the treatment.

We will start with reporting estimates using only our stratification design. Given the school-level pairing of our design into stratas, we first aggregate the student-level data to schools level and estimate the intention-to-treat effect as

$$\tau = \frac{1}{G}\sum_{g=1}^{G}\tau_g \tag{1}$$

where G is the number of strata and τ_g is the within strata difference in average outcome between the treatment and control groups (see Athey and Imbens (2017), section 6.2 for discussion). Our baseline estimator for the variance of this estimator is

$$\hat{\mathbb{V}}(\tau) = \frac{1}{G(G-1)} \sum_{g=1}^{G} \left(\hat{\tau}_g - \hat{\tau}\right)^2 \tag{2}$$

which we will use for constructing confidence intervals. However, we will assess the statistical significance of the treatment effects using Fisher exact p-values (see Athey and Imbens 2016, section 4.1, for discussion).

While school-level analysis provides a natural starting point, we expect to be able to increase precision by conditioning on a rich set of background characteristics. As discussed above, we do not have a full knowledge on which variables will be available at the time of implementing our analysis, but we expect to use information along the following dimensions:

- Parental characteristics such as education, country of birth, income and employment.
- Sibling characteristics (similar to register-based information we will use for students in the treatment and control groups).
- School characteristics. These include the average parental characteristics at the school level; past GPA of the school; past upper-secondary enrollment level of the school.
- Neighborhood characteristics. These include average education, employment level of the neighborhood.

Finally, there are two versions of the intervention (light vs. intensive treatment) and thus we will report results comparing

1. All students in treatment vs. control schools
2. Students assinged to receive intensive treatment vs. no treatment
3. Students assinged to receive light treatment vs. no treatment
4. Students assinged to receive intensive treatment vs. light treatment

5.2 Accounting for multiple inference

As we report results on many outcomes, we discuss the significance of our estimates both in isolation and as a member of a family of hypotheses. Our outcomes can be classified into three families: education outcomes, application outcomes, and socio-economic outcomes. To fix the familywise error rates, we follow Kling et al. (2004) (see also Finkelstein et al. (2010)). We follow Kling et al. (2004) in appying several different adjustments to the p-values that are the smallest familywise significance levels at which one can reject the null hypothesis that there are no effects on any of the members of the outcomes in the family.

5.3 Heterogeneous effects

We will examine treatment effect heterogeneity by gender, grades, immigration status and family background.

References

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

Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs, NJ 1986*.

Finkelstein, A., S. Taubman, H. Allen, J. Gruber, J. Newhouse, B. Wright, and K. Baicker (2010). The short-run impact of extending public health insurance to low-income adults: Evidence from the first year of the oregon medicaid experiment. *Analysis plan*.

Goux, D., M. Gurgand, and E. Maurin (2017). Adjusting your dreams? high school plans and dropout behaviour. *The Economic Journal 127*(602), 1025–1046.

Kling, J. R., J. B. Liebman, et al. (2004). *Experimental analysis of neighborhood effects on youth*. John F. Kennedy School of Government, Harvard University.

Meichenbaum, D. (2017). Stress inoculation training: A preventative and treatment approach. In *The Evolution of Cognitive Behavior Therapy*, pp. 117–140. Routledge.

Salmela-Aro, K., P. Mutanen, P. Koivisto, and J. Vuori (2010). Adolescents' future education-related personal goals, concerns, and internal motivation during the towards working life group intervention. *European journal of developmental psychology 7*(4), 445–462.

Vuori, J., P. Koivisto, P. Mutanen, M. Jokisaari, and K. Salmela-Aro (2008). Towards working life: Effects of an intervention on mental health and transition to post-basic education. *Journal of Vocational Behavior 72*(1), 67–80.