Applied Microeconometrics I
Aalto University, Department of Economics

**Final exam — Answer Key**
*October 12th, 2026*

*You have 3 hours to take the exam, though the exam is not intended to take the whole period. You should be able to answer each question with a graph, number, and/or no more than three concise sentences. None of the problems require a calculator. Enjoy and have fun!*

1. Define the fundamental problem of causal inference in one sentence. (1 point)

   *Answer: The fundamental problem of causal inference is that we can never simultaneously observe the potential outcome of the same economic unit under both treatment and control; we only observe one state of the world for each unit.*

2. In 2024 the Helsinki metropolitan area decided to try to boost the productivity of the startup community in the region – and wants to do so in an evidence-based way. They randomly assign a subset of all technology firms employing 5 or less people in 2024 and established between April 1st 2021 and April 1st 2023 the opportunity to access a 150K euro hiring benefit if these firms expand hiring. In total 186 firms meet these criteria, and they randomly assign 62 firms access to the hiring credit. Of the 62 firms with access to the credit, 31 use it to hire more employees.

   In 2026, firms using the credit employ an average of 6.2 workers while those in the comparison group employ an average of 4.2 workers. Firms in the treatment group which do not access the credit employ an average of 3.2 workers.

   (a) What is the effect of access to hiring credits on firm size? (1 point)
   *Answer: The intention-to-treat (ITT) effect is 0.5 workers (*$6.2 \times 0.5 + 3.2 \times 0.5 = 4.7$ *average workers in the treatment group minus* $4.2$ *in the control group).*

   (b) What is the impact on employment for firms which actually use the credit? (1 point)
   *Answer: Using the Wald estimator (scaling the ITT by the first-stage compliance rate of* $31/62 = 0.5$*), the local average treatment effect is 1.0 worker (*$0.5/0.5$*).*

   (c) You are asked to advise policy-makers regarding the program's impact. Why might you prefer to emphasize the result in (a)? Why might you prefer to emphasize the result in (b)? (1 point)
   *Answer: You would emphasize result (a) (ITT) if policy-makers care about what offering the hiring credit to a similar set of firms might do (taking into account the fact that not all firms will use it). You would emphasize result (b) (LATE) if they want to know the direct causal effect on firms that actively choose to participate: i.e. something like how much does relaxing resource constraints for hiring affect firm productivity.*

   (d) Define the terms *compliers, always-takers, never-takers, and defiers* in this context. (1 point)

*Answer: The 31 who use the credit are made up of both compliers who use the credit if and only if granted access and always-takers use it regardless of access (this always-takers is a bit harder to conceive of here in terms of using credits, but you can imagine hiring with or without the program); never-takers never use it regardless of access (presumably the 31 firms which do not use it but were given access are never takers); defiers only don't hire if they get access to the hiring credit – or use it only if denied access (if we assume defiers don't exist, we can identify the never-takers in the treated group).*

3. It turns out you have access to revenue data on all firms in the sample going back to April 1st 2021. You want to use a differences-in-differences strategy to estimate the effects of using the hiring credits comparing the 31 firms in the treatment group which use the credit to the 31 firms also in the treatment group which do not.

   (a) What is the identifying assumption you would need for this strategy to work? (1 point)
   *Answer: Parallel trends (in potential outcomes): In the absence of using the credit, the revenue trajectories of the firms that used the credit and those that did not would have followed parallel paths over time.*

   (b) Write down the econometric specification you would use. (1 point)
   *Answer:* $Y_{ft} = \beta D_{ft} + \alpha_f + \lambda_t + \epsilon_{ft}$*, where* $\alpha_f$ *and* $\lambda_t$ *are firm and time fixed effects, and* $D_{ft}$ *is an indicator for active credit use.*

   (c) What could you do to probe for the validity of your design? (1 point)
   *Answer: Test for parallel pre-trends by plotting the differences in user vs. not user firms each year (event-studies), and run a covariate balance check using the exact same specification as in the DiD, but just switching the outcome with some covariates which should not be affected by the policy (e.g. CEO age).*

   (d) Why might the tests you use in (c) fail in this setting? (1 point)
   *Answer: Endogenous treatment timing or anticipatory shocks can cause differential pre-trends between users and non-users, violating parallel trends. In this setting in particular, we might be concerned that firms which use the treatment specifically chose to use it – given the timing it was offered: this may mean that even if the trends were relatively similar prior to the treatment, they may have just hit a point in their startups where they were starting to grow, etc... and that's why they actually used the program. Put simply, they used the program when offered treatment specifically because they were just about to enter a different growth phase (anyway). NOTE, a wrong answer would say that the groups are different in observables, or mention selection bias without mentioning that the selection would have to do with the timing (if parallel trends were OK prior to treatment): i.e. differences in levels is not a problm for DiD.*

4. Helsinki also gives you access to data on all other firms in the region all the way from 2015. You want to practice all your econometric tools, and decide to use a regression-discontinuity design to estimate the effects of the program.

(a) You focus on the aspect of the assignment rule wherein, conditional on having been established in the right period, firm size determines participation in the randomization. What is the econometric specification you would use to estimate the reduced form effects of the program? (1 point)

*Answer:* $Y_i = \alpha + \beta \mathbb{D}_i + f(X_i - c|D_i = 0) + f(X_i - c|D_i = 1) + \epsilon_i$*, where* $X_i$ *is firm size (the running variable),* $c$ *is the cutoff, and* $f(\cdot)$ *are functions of the running variable, allowed to vary on either side of the cutoff.*

(b) What is the identifying assumption in this design? (1 point)

*Answer: The identifying assumption for RDD is that the potential outcomes would have evolved smoothly across the threshold in the absence of the treatment.*

(c) Name two things you could do to test for the validity of the design. (1 point).

*Answer: (1) Perform a McCrary density test to check for manipulation of the running variable. This tests for the bunching of observations near the cutoff. (2) Check for covariate balance at the threshold.*

(d) Calculate the first-stage estimate on credit use yielded by the regression discontinuity design? (give me a number) (1 point)

*Answer: 0.16 (or* $16\%$*), representing the fact that only half the firms assigned treatment actually use the credits, and only 1/3 of all eligible firms were assigned treatment– assuming this is uniformly distributed within the eligible firms. However, if larger firms hire more/less people, the first stage could be higher/lower.*

(e) You use the first-stage estimate to scale the reduced form estimates on revenue. Why might these estimates of the use of hiring credit use differ from those on using the hiring credit from the designs described in Questions 2 and 3? (1 point)

*Answer: The RDD estimates a Local Average Treatment Effect (LATE) specifically for firms at the size threshold, whereas Question 2 estimates LATE for the experimental sample and Question 3 estimates ATT with potential selection bias. If the probability of using the credit is uniform across the size distribution, then the experimental and RDD late could be the same, but they don't have to be.*

(f) Name an alternative regression-discontinuity strategy you could have used? (1 point)

*Answer: A regression discontinuity design using the establishment date cutoff (between April 1st 2021 and April 1st 2023).*

5. You are curious about whether or not firms which use the hiring credits end up hiring not just more workers but more productive workers.

   Workers in the tech-industry experience high-levels of mobility from one firm to another, particularly in startups. You have access to a linked employer-employee data set which contains (only) worker identifiers, measures of wages, and also firm identifiers all the way from 2015 onwards.

   You decide to estimate worker productivity from the wages of the startup employees at their prior employers using the following regression equation: $\ln w_{it} = \alpha_i + \psi_j + \epsilon_{it}$, where $w$ measures wages, $i$ indicates worker, $t$ indicates year, and $j$ indicates firm.

(a) What is the advantage of this way of estimating worker productivity have from simply looking at the wages of workers at their prior firm? (1 point)

*Answer: The two-way fixed effects model separates portable worker productivity ($\alpha_i$) from firm-specific pay premiums ($\psi_j$), avoiding confounding firm wage policies with pure worker productivity.*

(b) Why might this approach fail in this context? (1 point)

*Answer: Violation of strict exogeneity due to endogenous worker-firm sorting based on time-varying match-specific shocks ($\mu_{ij}$) or limited worker mobility across firms (firms not strongly enough connected via workers who switch firms).*

6. It is expensive for the cities to keep all the money available for firms if firms are unlikely to use it. Prior to randomization, the government collected detailed survey information on firms, their needs, their aims, and their constraints. You have an idea that you could use machine learning to predict which firms actually might end up using hiring credits in the future.

(a) What could you do to avoid "overfitting" when you predict which types of treated firms actually use the credits? (1 point)

*Answer: There are many possible ways to use ML to avoid overfitting. Regularization methods like LASSO and cross-validation methods across folds of the data can be helpful. Critically, it's important to gauge fit in a training subset and only evaluate predictive performance on a strictly held-out test dataset.*