Report on **Bachelor / Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Jakub Komárek**
Advisor:	**Václav Korbel**
Title of the thesis:	**Estimating the Effect of Teaching Assistant Support on Educational Outcomes**

OVERALL ASSESSMENT *(provided in English, Czech, or Slovak):*

Please provide a short summary of the thesis, your assessment of each of the four key categories, and an overall evaluation and suggested questions for the discussion. The minimum length of the report is 300 words.

Short summary

The goal of this thesis is to estimate the effect of Teaching Assistants (TAs) on pupil's education outcomes in Czech primary schools. The author provides intuitive discussion about potential effects of TAs and based on them formulates three hypotheses: (1) the more TAs per student the better average educational outcomes, (2) School characteristics have an effect of TAs usage, and (3) Schools closer to an Education Assesment Facilitiy (EAF) have more TAs per student. Data on the level of schools are used to test these hypotheses. As education outcomes the author uses grade repetition rate and school non-completion rate. Even though advanced identification methods are used (propoensity score matching and instrumental variable estimation) unreliable, inconclusive results are obtained, what is acknowledged in the thesis.

Contribution

The questions asked in this master thesis are extremely relevant and its results could constitute an important input for policy makers. The author did not manage to obtain informative results while working on this thesis, which, I believe, is mainly due to limited amount of time.

Methods

Identification (i.e. unbiased estimation of the causal effect) of the effect of teaching assistants on students education outcomes is very difficult and the author of this thesis mentions some of the difficulties. First, there is a problem of reverse causality. TAs are usually assigned to 'problematic', lower-performing students, which might result in a negative corelation between presente of a TA and students' performance. Such correlation is driven by the causality from student to TA, while this thesis aims at estimating the effect from TA to student. Another difficulty is that this thesis works with aggregate data on school level and so only an aggregate effect might be estimated. Earlier studies found that TAs have a positive effect on disadvantaged students (the ones they are assigned to), but the total effect on the classroom or the whole school might be minimal.

The author deals with these issues by (1) applying matching estimation and instrumental variable estimation, (2) zooming on outcomes that are driven mainly by ‚problematic' students: grade repetition rate and school non-completion rate.
I appreciate the effort to properly identify the effect in question, even though the final application and description of methods listed in (1) above has quite a few weak points.

The author proposes to break endogeneity of the variable measuring TAs presence by using distance from school to the closest Education Assesment Facilitiy (EAF) as an instrument. The argument (which is properly outlined in the thesis) is that distsnce between school and EAF is exogenous (i.e. not correlated with school outcomes) and that with high distance to EAF parents are less likely to visit such a facility and get an official confirmatoin that their child needs a TA.This all makes sense in theory, but in practice it appears that in the data the correlation between distacne to the nearest EAF is positively associated with the number of TAs per students (i.e. the further the EAF, the more TAs). The author hypothesizes that this might be because there are more disadvantaged students in rural

Report on **Bachelor / Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Jakub Komárek**
Advisor:	**Václav Korbel**
Title of the thesis:	**Estimating the Effect of Teaching Assistant Support on Educational Outcomes**

areas that are far from EAFs. But even if just some of them get a diagnosis and are eligible for a TA, this still results in sligthly more TAs per pupil than in schools located in urban areas. Another possible explanation given in the thesis is that staff of EAFs sometimes travels to schools with many disadvantaged students. These issues could be partially dealt with by controling for urban/rural location of the school (not done in the thesis), by controling for past share of disadvantaged students, and/or by removing Prague from the estimation sample (not done in the thesis).
Given that the 1st stage of the IV 2SLS estimation does not work as expacted, the author does not proceed with this estimation approach.

As an alternative identification strategy the author proposes propensity score matching. Let me stress here that, in contrast to what the author claims, **matching does not solve endogeneity issues better than OLS**! Matching may be used when we believe that selection to treatment is based on observables. When we include all such observables in an OLS model, there would be nothing left in the error term that is correlated with the treatment variable (with the variable measuring the presence of TAs). This means that matching is not better than OLS in dealing with endogeneity. It seems that the author of this thesis is not aware of it. Matching improves over OLS in making sure that treated and nontreated observations are similar enough to compare them. In the specific case of this thesis, it might happen that schools with many TAs are regulairly different than schools with few TAs. This means that OLS has to extrapolate to estimate the effect of TAs. Matching avoids extrapolation and estimates the treatment effec using only such schols that are similar to each other.

When choosing the matching variables it is necessary to include all relevant information that could affect seletion to treatment. Here: that could affect that a school has many TAs. Matching is not well explained in the thesis, so I am not able to figure out if it was done correctly or not. For example, what has been done to the sample so that the „after matching' balance of propensity schores looks as in the bottom part of figure 6.2? Which matching algorithm was used, which weights are used? How does the author deal with observations out of the common support (with very low propensity schore)? It's good to know that matching resulted in balanced characteristics between treatment and control groups (Figure 6.4), but we do not knw what was done next.

The author just writes that matching estimation resulted in an insignificant estimate of the treatment effect. But we know nothing about HOW this estimate was obtained (nearest neighbor matching? Kernel metching? Another approach?) nor how it's statistical significance was computed (bootstrap?). Maybe the null effect is driven by the definition of the treatment, which is arbitrary? What if top 25% schools were conpared with bottom 25% (in terms of TA usage) as in Black et al. (2004)?

The baseline equation provided in the Methodology section relates CHANGE in school outcomes between the periods 2019-2022 and 2016-2019 to the CHANGE in the number of TAs. It is not clear at all why these two periods are compared. Given that growth in TAs was continuous over the whole period between 2016 and 2022, the division onto two periods (called by the author pre-treatment and post-treatment) seems arbitraty. There is not discussion about it. From the point of view of further applied identification strategies this before-after specification is useless and just introduces confusion.

Literature

The literature review is based on the literatue review summarized in Farrell, Alborz and Howes (2010) plus two newest studies (Andersen et al. 2020 and Hemelt et al. 2021). It quite well presented and highlights important results. What I miss is some evidence from the Czech Pepublic, even if it would not be based on a robust regression analysis.

Report on **Bachelor / Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Jakub Komárek**
Advisor:	**Václav Korbel**
Title of the thesis:	**Estimating the Effect of Teaching Assistant Support on Educational Outcomes**

Manuscript form

Manuscript form is the weakest part of htis thesis. It is written in quite a chaotic way with may typos, missing words, not connected paragraphs and confusing sentences. The introduction stresses inclusive education and writes a lot about students with special educational needs (SEN), so the reader has an impression that the thesis is about inclusive education. The role of TAs is opened on the 2nd page of the introduction and without any motivation the author writes that the goal of this thesis is to „enrich existing literature studying TAs effects". It's not clear in the intro whether effects on disadvantaged students or the effects on other students, or total effects are considered. And why. Methodoloogy is outlined in the introduction in an extremely chaotic way.

The author stresses in the thesis that year 2016 marks important changes in provision of TAs in Czech primary schools. It is not clear in the text how big and suddent change it was. We can see in Figure 2.1 that number of TAs (per teacher) was gradually and visibly increasing between 2016 and 2022, however we do not know how it looked before 2016. This figure, however, does not support further applied methodology.

Methodology and results sections also look chaotic and are very difficult to read. It looks like a set of paragraphs that do not necessairly connect to each other and the reader is left to complete open thoughts of the author. For example, the last paragraph before section 6.1. cites findings of Munich and Protivinsky (2022), but does not explain why and what does it mean for the analysis presented in this thesis. Table 6.1 is redundant (Figure 6.4. provides the same information in much better way). ESS is discussed in the first paragraph on p. 31 and we do not know what does it mean for the analysis, etc.

Results are discussed relatively well, even though there are some issues with them (as discussed in the previous section). I am just confused with Figure 7.1 which reporth positvive effects of number of disadvantaged students on TA presence, while Table 7.2 reports negative coefficients.

The discussion section is relevant and explains what might have went wrong and how potentially one could improve it.

There is a long Policy recommendations section which, in my opinion, should not be a part of a thesis. Mainly, because there are actually no results on which recommendations could be made. Thus, recommendations are made based on… author's believes? Expectations? They are pretty critical of the current systém even though the thesis started with a statement „The purpose of this thesis is not to advocate for or against inclusive schooling but to provide evidence for evidence-informed school policymaking"

Overall evaluation and suggested questions for the discussion during the defense

The thesis asks a very relevant question, provides decent literature review and theoretical/intuitive discussion that motivates the analysis. The author proposes solid identification methods (though reliance on matching is too optimistic), but it turns out that the proposed instrument does not work as expected and matching gives a null result. The whole thesis is written in a very chaotic, hard to understand way.

Report on **Bachelor / Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Jakub Komárek**
Advisor:	**Václav Korbel**
Title of the thesis:	**Estimating the Effect of Teaching Assistant Support on Educational Outcomes**

In my view, the thesis fulfills the requirements for a bachelor thesis at IES, Faculty of Social Sciences, Charles University, I recommend it for the defense and suggest a grade C.
The results of the Turnitin analysis do not indicate significant text similarity with other available sources.

SUMMARY OF POINTS AWARDED *(for details, see below):*

CATEGORY		**POINTS**
Contribution	*(max. 30 points)*	27
Methods	*(max. 30 points)*	23
Literature	*(max. 20 points)*	18
Manuscript Form	*(max. 20 points)*	12
TOTAL POINTS	*(max. 100 points)*	**70**
GRADE	**(A – B – C – D – E – F)**	**C**

NAME OF THE REFEREE: Barbara Pertold-Gebicka

DATE OF EVALUATION: 15.6.2023

Digitálně podepsáno (15.6.2023)
Barbara Pertold-Gebicka

_____

Referee Signature

EXPLANATION OF CATEGORIES AND SCALE:

CONTRIBUTION: *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

METHODS: *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

LITERATURE REVIEW: *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

MANUSCRIPT FORM: *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

Overall grading:

TOTAL	**GRADE**
91 – 100	**A**
81 - 90	**B**
71 - 80	**C**
61 – 70	**D**
51 – 60	**E**
0 – 50	**F**