Economics Department Discussion Papers Series

ISSN 1473 – 3307

Does formative feedback help or hinder students? An empirical investigation

Carlos Cortinhas

Paper number 17/01

Does formative feedback help or hinder students? An empirical investigation

Carlos Cortinhas*

Abstract

The link between formative assessment and student performance is not entirely clear in the existing literature with some previous studies showing contradicting results. Although the debate is very old (since mid-1800s), previous research is almost exclusively based on elementary and secondary school students. This paper attempts to add to the existing literature by focusing on data from a large scale experiment (a class of 578 UK first year undergraduate students enrolled on Introduction to Statistics) to determine whether online, formative (non-compulsory) homework helps or hinders students. The results suggest that completing formative assessment tasks contributes to higher grades but only for good students. The result is robust to a variety of specifications and after controlling for a large number of student characteristics (including nationality, gender, ethnicity, whether a student has completed a Maths or Economics A-Level in Secondary School, amongst others) and the level of student ability/effort. This study shows, therefore, that formative homework might contribute to amplifying inequalities amongst students and that other strategies are needed to close the gap between the top and bottom performing students.

Keywords: homework, online learning, formative assessment.

JEL Classifications: C21; I21; I24;

1. Introduction

The debate over the merits of formative assessment (and homework in particular) and its effects on student attainment is an old one and the issues raised are far from being resolved. Gill and Schlossman (2004) provide a brief history of homework in America and report that the controversy of whether homework improves performance started in the late 1800's with 'a crusade against homework', and the debate has continued ever since with attitudes to homework being cyclical over time.

* The University of Exeter, UK and The Economic Policies Research Unit (NIPE), The University of Minho, Portugal; c.cortinhas@exeter.ac.uk. The usual disclaimer applies.

The debate is almost exclusively based on elementary and secondary school students and comes mostly from the United States. Recently, the debate has opened up to include higher education and to attempting to determine the effectiveness of different kinds of homework in a variety of different contexts.

Although difficult to summarize given the extent of the literature on the subject, the key findings of the literature seem to be as follows[1].

To date, the link between homework and student achievement is not entirely clear. Some studies show positive effects (e.g. Keith, Diamond-Hallam and Fine (2004), Trautwein et al. (2009)) while others showed no or negligible effects (see for example Cooper (2009) for a survey)[2]. Interestingly, there is a strong generalized agreement that 'too much' homework and homework that is set 'too often' is harmful but there is no agreement on what 'too much' and 'too often' is. There are significant differences between the effect of homework across countries, and, further differences are evident as regards such variables as age, gender, ethnicity amongst others. Girls seem to benefit more than boys[3] (Falch and Ronning (2012)), for example; homework seems to be more beneficial for older students (see for example Cooper 1989), and homework might be beneficial in some subjects (Maths) but not in others (Science, English and History) (Eren and Henderson (2011)). Additionally, some ethnic groups might benefit more than others (Keith and Benson (1992), and students with learning disabilities benefit from homework but only under certain conditions (Cooper and Nye (1994)).

A number of recent papers have particular relevance to the data, methodology and results of this paper. Ronning (2011) found that for elementary school students homework amplifies existing inequalities as students from low-income homes do not benefit as much from homework as students from high-income homes[4].

Torst and Salehi-Isfahani (2012) found that the completion of individual online homework assignments early in a Principles of Economics course appeared to

---

[1] For a survey of the literature on homework as well as a description of the suggested pros and cons of homework see for example Cooper (1989) or Cooper, Robinson and Patall (2006).

[2] It is important to note, however, that the subset of the literature that studies the relationship between *time* spent on homework and achievement is mostly positive. This does not, however, "fully reflect the complex relationship between homework and achievement" Ulrich et all (2009), p.77.

[3] Falch and Ronning (2012) find that the effect of the number of homework tasks on student achievement is positive in most countries but wide variations exist. Interestingly, in the case of one country (Sweden) a significant (at the 5% level) negative relationship between homework and attainment was found.

[4] It is important to note, however, that we would expect this effect to be larger for elementary school students than university students as this effect is likely to also capture lower engagement levels from parents in the case of low-income students and this is less relevant for higher education.

modestly improve student performance on related questions in the mid-term exam but also, perhaps paradoxically, found that missing one homework assignment did not negatively affect final exam performance.

Self (2013) found that performance on compulsory online homework assignments in a Principles of Macroeconomics class did not impact on test grades but also that (voluntary) access to the website to practice additional problems did impact positively on student results. This result was found to be true even for weaker students (students with lower GPAs) although the study could not determine whether the key to this success was the use of the online resource or whether the result is simply explained by the greater effort employed in doing this extra work. Grodner and Rupp (2013) on the other hand, using a field experiment where students were randomly assigned into homework-required and not-required groups in a Principles of Economics class, found that homework submission had a large positive effect on test performance and on retention rates.

Rhodes and Sarbaum (2015) found that allowing students to have multiple attempts when completing their online homework online tasks leads to guessing behaviour and grade inflation, hence improving score without improving learning outcomes.

Lee, Courtney and Ballasi (2010) found no statistically significant difference in student performance between the use of an online homework tool (Aplia) and traditional homework methods. Interestingly, however, when they focused on the subset of students with the highest grades (students that received either an A or B in the Principles of Microeconomics course), the group doing using the online homework tool improved their scores by nearly two points over the group of students who used traditional homework methods.

Eren and Henderson (2008) using both parametric and non-parametric techniques found hours of homework assigned to be an important determinant of student achievement but not for all students. Their study, which focused on eighth grade students, found homework to be effective in improving performance for high and low achievers but not for the average achievement group.

Finally, Dahlgran (2008) found online homework to increase learning through increased student study-time allocations, i.e., online homework might increase the overall time students allocate to studying a topic and as a result lead to an improvement in their performance.

This paper attempts to add to the existing literature by focusing on data from a large scale experiment to determine whether online, formative homework helps or hinders students in the context of higher education.

2. Data and Empirical Methodology

The data for this study is composed of all students registered to take the Introduction to Statistics (undergraduate first year) module in the academic year 2014/15 at the University of Exeter. This module is compulsory for all Business School students and is open (as an elective) to all university students and is comprised of 579 students of 50 different nationalities registered in 16 different undergraduate programmes.

The homework was part of the formative assessment (non-compulsory) and as such did not count towards the student's final grade. Students were asked to complete a total of 10 homework tasks over a period of 10 weeks on the module's virtual learning environment page. The tasks consisted mostly of multiple choice questions, some of which involved basic calculations before answering. Students were allowed to have as many attempts as they needed to answer, but to avoid the guessing behaviour described by Rhodes and Sarbaum (2015), each new attempt led to a different set of random questions being generated as giving students individually unique exercises as opposed to exercises that are identical to all students has proven effective in previous studies (see for example Joerding (2010)). Finally, in order to encourage students to adopt patterns of regular study, students were given a window of about 10 days to complete each of the homework tasks, after which the homework task became unavailable.

The data for this study includes data on student performance for summative assessment (final exam grade), formative assessment (average homework grade) and data on a large number of student observable traits, including age, gender, ethnicity, country of domicile, undergraduate degree, whether the student has a disability, whether the student came from a state-funded or private-funded independent school and whether the student took an A-level in Maths and/or Economics at high school[5]. Finally, the weighted average grade the student achieved across all year 1 modules (not including their grade in Introduction to Statistics) is also included as a proxy for (unobservable) ability and effort.

Table 1 presents the summary statistics of the data used in this study.

Table 1: Sample Statistics of Key Variables

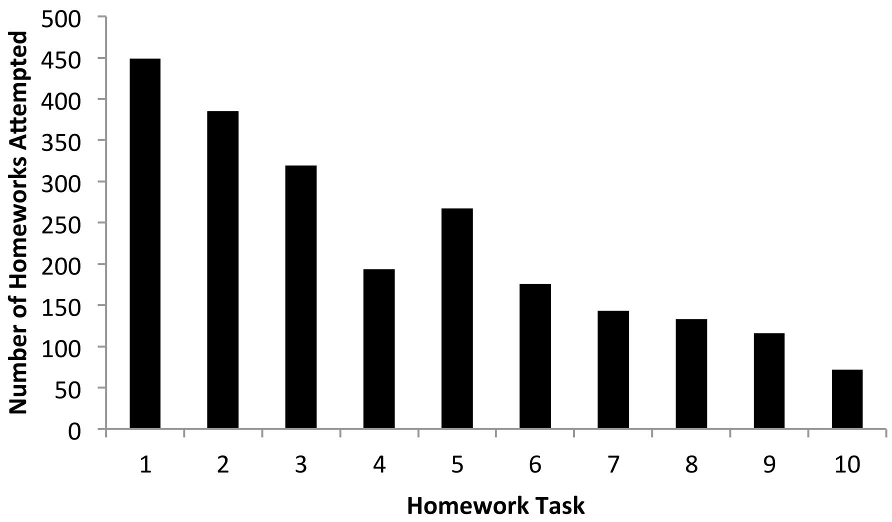
Variables	Mean	Standard Deviation
Exam Grade	56.628	19.314
Homework grade	30.947	24.707
Ability/Effort	63.105	11.796
Maths A-level	0.808	0.395
Economics. A-Level	0.887	0.317
Gender (male=1)	0.614	0.487
Disabled student	0.052	0.222
Independent School	0.316	0.465
State School	0.641	0.480
Age	20.25	1.578
Ethnicity		
White	0.689	0.473
Black	0.029	0.169
Chinese	0.143	0.350
Indian	0.036	0.187
Pakistani	0.007	0.083
Other Asian	0.064	0.245
Other	0.031	0.173
UG Programme of study		
Accounting and Finance	0.187	0.390
Business and Accounting	0.029	0.168
Business and Economics	0.135	0.341
Economics	0.349	0.477
Economics and Finance	0.162	0.396
Economics and Econometrics	0.016	0.124
Accounting	0.067	0.251
Flexible Combined Honours	0.022	0.148
Other	0.033	0.178
Country of domicile		
England	0.603	0.489
China	0.067	0.250
Hong Kong	0.040	0.195
Singapore	0.031	0.173
Wales	0.043	0.203

Table 1 provides some interesting facts about this class. First, students performed worse on average in Statistics (57%) than in their other subjects, as measured by the ability proxy (63%) which seems to confirm the view that students find the subject challenging. Most students were male (61%), had attained a Maths-A Level (81%) and Economics A-Level (89%). 5% of the students were disabled and most

[5] In the U.K., independent schools are also paradoxically known as public schools. The A-level is a high school qualification offered to students completing secondary or pre-university education. Students are required to study at least 3 A-levels (or equivalent) in order to be accepted to study for an undergraduate degree. The number of nationalities in the sample is over 50.

came from State Secondary Schools (64%). About 60% of students were English and most were taking an undergraduate degree in Economics (35%). The average exam grade was 56% and the standard deviation was quite large (almost 20%) whilst the average Homework grade was only 31%, suggesting that the uptake in the homework tasks was on average quite low[6]. In fact, and as shown in Figure 1, the number of homework tasks completed by week dropped steadily throughout the term, and went from around 450 students completing their tasks in week 1 to just over 70 students doing so in week 10[7].

Figure 1: Average number of homework tasks completed by week



We assume that the impact of homework on the exam score (EG) of pupil *i* can be explained by the following simple education production function:

$$EG_i = \beta HW_i + \gamma AB_i + \delta x_i' + \varepsilon_i \quad (1)$$

Where EG_i is the final exam grade for student I, HW_i is the average grade of 10 homework tasks completed by student i, AB_i is a proxy for (the unobserved) student's i level of ability/effort, x_i' is a vector of observed student traits (including Nationality, Age, Gender, Degree, whether the student has an A-level in Maths and/or Economics, Ethnicity, etc.) and ε_i is a zero mean, normally distributed error term.

[6] The homework grade was computed as $Homework\ Grade_i = \frac{\sum_{j=1}^{10}(Homework\ Grade_j)}{10}$ so that if a student missed a task, the average was heavily penalised.

[7] Week 6 of term 1 is a reading week for first year students (where students have no classes) and this explains the slight increase in completions when compared with the week before.

The variable AB_i was added to control for the (unobserved) different levels of ability and effort as otherwise it would be impossible to account for the separate effect of the completion of homework on final exam performance. This variable was computed by calculating the weighted average of each student's year grades in all modules taken during that year of study (excluding the grade for Introduction to Statistics).

3. Results

Table 2 presents the results of the baseline model, in five alternative regressions with alternative combinations of the vector x_i' variables[8].

Table 2: Regression Results full sample

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Exam Grade	Exam Grade	Exam Grade	Exam Grade	Exam Grade
Homework grade	0.0649**	0.0741**	0.0973**	0.1102**	0.1116**
	(0.0152)	(0.0183)	(0.0149)	(0.0171)	(0.0204)
Ability	0.9883**	1.0476**	0.9954**	0.9909**	1.0227**
	(0.0630)	(0.0728)	(0.0481)	(0.0645)	(0.0604)
Maths A-level		-7.1529**	-2.9755*		-7.5490**
		(1.1739)	(1.2910)		(0.6520)
Economics. A-Level		3.7017		-2.4051	4.9496**
		(2.6541)		(1.1804)	(1.2408)
Constant	-8.8274*	-11.9819	4.1036	-19.6383	-9.7809
	(3.9116)	(22.6751)	(12.2662)	(15.6561)	(15.3912)
Observations	577	222	302	283	222
R-squared	0.389	0.417	0.480	0.483	0.528
RMSE	15.13	14.64	14.84	14.19	13.81
LogLikelihood	-2384	-907.3	-1233	-1142	-884.0
Other controls:					
Gender	Y	Y	Y	Y	Y
Age	N	Y	N	N	Y
Programme of study	N	N	Y	Y	Y
Ethnicity	N	N	Y	N	Y
Disability	N	N	Y	Y	Y
Country of domicile	N	N	N	Y	Y
State School	N	N	Y	Y	Y

Heteroscedasticity-robust standard errors in parentheses, clustered by country of domicile
** p<0.01, * p<0.05

[8] The STATA option cluster was used so that the observations were clustered into countries of domicile to account for the possibility that the observations are correlated within countries (but are expected to be independent across countries). Cameron and Miller (2015) for example, state that "failure to control for within -cluster error correlation can lead to very misleadingly small standard errors, and consequent misleadingly narrow confidence intervals, large t-statistics and low p-values". The cluster option affects the standard errors and variance-covariance matrix of the estimators but not the estimated coefficients.

All regressions in this section were estimated to account for the possibility that the error term contains clustered errors, i.e. that observations within countries of domicile are correlated in some unknown way, inducing correlation in ε_i within those groups of students from the same country of domicile but that different countries of domicile do not have correlated errors. The possibility of observations being correlated within a country of domicile but uncorrelated across countries seems reasonable as all students will experience different education systems and cultural values according to the country of origin.

Several results immediately stand out. First, the coefficient for ability is highly significant and yields a result around 1 in all specifications, suggesting that ability and effort is a strong determinant of academic performance. Secondly, the coefficient for homework grade is also significant at the 0.01 level and has the expected positive sign. This suggests that completing and performing well on formative homework tasks positively contributes to a better final exam grade. This is particularly significant as the homework tasks were constructed to allow students to have multiple attempts and thus learn with the activity. Finally, the Maths A-level coefficient is significant in some specifications but has an unexpected negative sign. One reason for that result might be that students with a Maths A-level, having covered some of the initial Introduction to Statistics content in the previous year, do not fully engage with the module as they believe they already know the content and end up performing worse that colleagues that have no background in statistics and therefore need to work harder.

Because the sample is fairly large we can also run a separate regression for the students who passed (i.e., achieved a pass mark of at least 40% in the final exam). Table 3 presents the same 5 specifications as Table 2 but uses as the sample only the students that passed the final exam.

The results are clearly generally in line with those in Table 2, with the major difference being the size of the coefficient for the variable Ability that drops from around 1 to a range between 0.69 and 0.83. Importantly, the coefficient for Homework Grade remains highly significant and positive in all five specifications.

Table 3: Regression Results with sample of students that passed the module

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Exam Grade	Exam Grade	Exam Grade	Exam Grade	Exam Grade
Homework grade	0.0506*	0.0951**	0.1064**	0.1088**	0.1281**
	(0.0196)	(0.0163)	(0.0287)	(0.0167)	(0.0170)
Ability	0.7219**	0.7648**	0.6931**	0.8314**	0.7854**
	(0.0757)	(0.0806)	(0.0625)	(0.0747)	(0.0891)
Maths A-level		-4.1943*	-1.4382		-3.8657*
		(1.4350)	(1.0711)		(1.3504)
Economics. A-Level		2.9437		-0.1834	2.8821*
		(1.9011)		(0.4654)	(1.2091)
Constant	13.2708**	-8.3462	15.5625	-13.0421*	-19.8980**
	(4.7012)	(11.9121)	(13.2711)	(3.9264)	(2.8327)
Observations	457	189	252	230	189
R-squared	0.281	0.304	0.412	0.447	0.490
RMSE	11.85	11.34	11.03	10.51	10.27
LogLikelihood	-1776	-723.6	-952.0	-857.0	-694.4

Heteroscedasticity-robust standard errors in parentheses, clustered by country of domicile
** p<0.01, * p<0.05

Similarly, Table 4 presents the results for the remaining students, i.e. only students that failed the module (i.e., students who did not achieve a pass mark of 40%).

Table 4: Regression Results with sample of students that failed the module

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Exam Grade	Exam Grade	Exam Grade	Exam Grade	Exam Grade
Homework grade	0.0380	-0.0901	-0.0154	0.0148	-0.0311
	(0.0213)	(0.0563)	(0.0421)	(0.0349)	(0.0447)
Ability	0.2268**	0.4563	0.4234**	0.1548	0.4241**
	(0.0726)	(0.2183)	(0.1299)	(0.1891)	(0.0114)
Maths A-level		-10.6888**	7.2192		6.4004
		(1.5495)	(3.6237)		(4.6000)
Economics. A-Level		9.3698		1.7988	0.7456
		(4.1009)		(1.3806)	(4.7350)
Constant	17.4030**	124.5005	-21.8831	43.8167	47.1250*
	(3.0507)	(86.6522)	(33.6799)	(51.6954)	(14.9152)
Observations	120	33	50	53	33
R-squared	0.119	0.250	0.472	0.305	0.764
RMSE	10.09	13.42	12.16	11.47	10.26
LogLikelihood	-445.6	-128.6	-185.5	-195.7	-109.5

Heteroscedasticity-robust standard errors in parentheses, clustered by country of domicile
** p<0.01, * p<0.05

The results show that for the sample of students that failed the module, the coefficient for homework grade becomes statistically insignificant all specifications.

The coefficient for Ability is only significant in three of the five specifications albeit it now yields a much lower coefficient (from 0.23 to 0.42).

4. Sensitivity Analysis

In this section, a series of tests is conducted to assess the robustness of the results presented in the previous section, focusing on whether the results are sensitive to the use of clusters and to alternative countries of domicile cluster definitions.

The decision of what to use as clusters, whether the number of clusters is large enough and whether clusters are of balanced size are important for the reliability of the results. In particular, it is now well established that clustering with a small number of clusters or with extremely unbalanced cluster sizes, may lead to incorrect inferences (see for example, Cameron and Miller (2015) for a technical discussion). That is, in some cases the cure (controlling for within cluster error correlation) might be worse than the disease (i.e. default standard errors can overstate estimator precision).

Therefore, the purpose of this section is to test whether the use of alternative strategies for the clustering of the errors significantly affects the results presented in Section 3.

First, equation (1) is estimated for the full sample with alternative clustering groupings. In this case, in an attempt to make clusters more balanced in size, clusters were mostly defined in terms of larger geographic regions[9]. The results are presented in Table 5[10].

When comparing to Table 2, it is clear that although the size of the standard errors varies slightly, the coefficients that were found statistically significant continue to be so at the same level of significance. This suggests that the results presented in Section 3 are not very sensitive to an alternative cluster definition for countries of domicile.

[9] The clusters were reduced to 13: Africa, Australasia, Central Asia, England, European Union, Middle East, North America, Northeast Asia, Other Europe, Other UK, Russian Federation, South Asia and South East Asia.
[10] The same patterns were found when estimating the model with the two smaller samples (students that passed and students that failed) presented in Tables 3 and 4. Those results are not presented here for the sake of brevity but can be provided upon request.

Table 5: Regression Results full sample with alternative country of domicile cluster groups

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Exam Grade	Exam Grade	Exam Grade	Exam Grade	Exam Grade
Homework grade	0.0649**	0.0741**	0.0973**	0.1102**	0.1116**
	(0.0201)	(0.0179)	(0.0140)	(0.0079)	(0.0128)
Ability	0.9883**	1.0476**	0.9954**	0.9909**	1.0227**
	(0.0648)	(0.0775)	(0.0560)	(0.0677)	(0.0667)
Maths A-level		-7.1529**	-2.9755*		-7.5490**
		(1.1882)	(1.2173)		(0.6088)
Economics. A-Level		3.7017		-2.4051	4.9496**
		(2.8111)		(1.3472)	(1.3991)
Constant	-8.8274*	-11.9819	4.1036	-19.6383	-9.7809
	(3.9743)	(26.3709)	(12.2662)	(16.6008)	(15.9934)
Observations	577	222	302	283	222
R-squared	0.389	0.417	0.480	0.483	0.528
RMSE	15.13	14.64	14.84	14.19	13.81
LogLikelihood	-2384	-907.3	-1233	-1142	-884.0

Heteroscedasticity-robust standard errors in parentheses, clustered by regions
** p<0.01, * p<0.05

Next, we provide in Table 6 an estimation of the baseline line model (full sample) without any clustering of the errors.

Table 6: Regression Results full sample without clustering of errors

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Exam Grade	Exam Grade	Exam Grade	Exam Grade	Exam Grade
Homework grade	0.0623**	0.0741	0.0973*	0.1102**	0.1116*
	(0.0268)	(0.0429)	(0.0386)	(0.0412)	(0.0449)
Ability	0.9900**	1.0476**	0.9954**	0.9909**	1.0227**
	(0.0626)	(0.1194)	(0.0920)	(0.1040)	(0.1303)
Maths A-level		-7.1529**	-2.9755*		-7.5490**
		(2.5959)	(2.3481)		(2.7160)
Economics. A-Level		3.7017		-2.4051	4.9496
		(4.0189)		(3.0628)	(4.2722)
Constant	-8.7206*	-11.9819	4.1036	-19.6383	-9.7809
	(3.9746)	(30.9612)	(27.2558)	(31.1698)	(35.4370)
Observations	578	222	302	283	222
R-squared	0.388	0.417	0.480	0.483	0.528
RMSE	15.15	14.64	14.84	14.19	13.81
LogLikelihood	-2389	-907.3	-1233	-1142	-884.0

Heteroscedasticity-robust standard errors in parentheses
** p<0.01, * p<0.05

The coefficients are the same but the standard errors are now larger in most cases, thus reducing the level of significance in some specifications (more specifically specifications (2), (3) and (5))[11]. The results continue to be, however, mostly in line with the results presented in Table 2 and suggest that the option to include cluster-robust standard errors was justified as the size of standard errors was found to be lower on those cases[12].

Finally, the use of degree programmes as clusters was also attempted. The rationale is that students taking the same degree are more similar that students in different degrees, independently of country of domicile. Table 7 presents the results.

Table 7: Regression Results full sample with programme of study as clusters

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Exam Grade	Exam Grade	Exam Grade	Exam Grade	Exam Grade
Homework grade	0.0623	0.0741	0.0973**	0.1102*	0.1116
	(0.0536)	(0.0481)	(0.0207)	(0.0502)	(0.0566)
Ability	0.9900**	1.0476**	0.9954**	0.9909**	1.0227**
	(0.0587)	(0.0582)	(0.0870)	(0.0550)	(0.0714)
Maths A-level		-7.1529*	-2.9755		-7.5490**
		(2.4257)	(2.5328)		(2.7160)
Economics. A-Level		3.7017		-2.4051	4.9496
		(4.0189)		(2.7597)	(3.5425)
Constant	-8.7206*	-11.9819	4.1036	-19.6383	-9.7809
	(3.6471)	(16.9041)	(15.8804)	(19.3517)	(17.4084)
Observations	578	222	302	283	222
R-squared	0.388	0.417	0.480	0.483	0.528
RMSE	15.15	14.64	14.84	14.19	13.81
LogLikelihood	-2389	-907.3	-1233	-1142	-884.0

Heteroscedasticity-robust standard errors in parentheses
** p<0.01, * p<0.05

The results are again in line with the results presented in Section 3 but again the standard errors were higher which reduced the number of statistically significant coefficients.

5. Final Remarks

This paper attempts to add to the existing literature on the impact of homework on student attainment by focusing on data from a large scale experiment in a university context to determine whether online, formative (non-compulsory) homework helps or hinders students.

[11] The coefficients are in fact very slightly different for the first specification. This is due to the fact that the information about country of domicile was missing for one student and therefore the sample size has increased by one observation.

[12] The fact that the cluster-robust standard errors are lower than the estimations without clusters can be explained by several factors, including noise (see for example Cameron and Miller (2015)) and seems to suggest that the model used in this paper does not have serious within-cluster correlation of standard errors.

The main result is that formative homework does help students in attaining better results but only for good students (i.e. students that managed to achieve at least a pass rate of 40%). For weaker students, formative homework appears to make no difference to their overall student attainment. This result is in line with the theoretical framework of Neilson (2005) that shows that in the presence of time-constraints, homework does not benefit all students and that in fact can increase the spread between the performance of the best and worst students. The results are, however, in contradiction to the results of Self (2013) that found that even weaker students benefit from formative online resources.

The results are robust to a variety of specifications and after controlling for a large number of student characteristics and the level of student ability/effort. This study suggests, therefore, that formative homework might contribute to amplify inequalities between students and that other educational strategies are needed to close the gap between the top and bottom performing students.

6. References

Cameron, A. & Miller, D. (2015). A Practitioner's Guide to Cluster-Robust Inference, *Journal of Human Resources*, Spring, Vol.50, No. 2, pp.317-373.

Cooper, H. (1989). *Homework.* White Plains, NY: Longman.

Cooper, H. and Nye, B. (1994). Homework for Students with Learning Disabilities: The Implications of Research for Policy and Practice, *Journal of Learning Disabilities*, 27(8), p. 470-479.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research, 76*(1), 1–62.

Dahlgran, R.A. (2008). Online Homework for Agricultural Economics Instruction: Frankenstein's Monster or Robo TA? *Journal of Agricultural and Applied Economics,* 40(1), 105-116.

De Witt, K & Van Klaveren, C. (2014). How Are Teachers Teaching? A Nonparametric Approach. *Education Economics*, 22(1), 3-23.

Eren, O. & Henderson, D. J. (2008). The impact of homework on student achievement. *Econometrics Journal,* 11, 326-348.

Eren, O. & Henderson, D. J. (2011). Are we wasting our children's time by giving them more homework? *Economics of Education Review,* 30 (5), 950-961.

Falch, T. and Ronning, M. (2012). Homework assignment and student achievement in OECD countries, Discussion Papers 711, Statistics Norway, Research Department.

Gordon, P. (1980). Homework: Origins and justifications. *Westminster Studies in Education*, 3, 27-46.

Grodner, A & Rupp, N. G. (2013). The Role of Homework in Student Learning Outcomes: Evidence from a Field Experiment. *Journal of Economic Education*, 44(2), 93-109.

Joerding, W. (2010). Teaching and Learning with Individually Unique Exercises. *Journal of Economic Education,* 41(2), 125-135.

Lee, W., Courtney, R.H. & BALASSI, S. J. Do Online Homework Tools Improve Student Results in Principles of Microeconomics Courses? *American Economic Review,* 100(2), 283-286.

Neilson, W. (2005). Homework and Performance for Time-Constrained Students. *Economics Bulletin,* 9(1), 1-6.

Nichols, A. & Schaffer, M. (2007). Clustered Errors in Stata. 2007 UK Stata Users Group meeting. Typescript available for download at http://www.stata.com/meeting/13uk/nichols_crse.pdf.

Parker, J. (2010). An Empirical Examination of the Roles of Ability and Gender in Collaborative Homework Assignments. *Journal of Economic Education,* 41(1), 15-30.

Rhodes, M. T. & Sarbaum, J. K. (2015). Online Homework Management Systems: Should We Allow Multiple Attempts? *American Economist,* 60(2), 120-131.

Ronning, M. (2011) Who Benefits from Homework Assignments?. *Economics of Education Review*, 30(1), 55-64.

Self, S. (2013) Utilizing Online Tools to Measure Effort: Does It Really Improve Student Outcome? *International Review of Economics Education*, 14, 36-45.

Trautwein, U., Schnyder, I., Niggli, A., Neumann, M., & Lüdtke, O. (2009). Chameleon effects in homework research: The homework–achievement association depends on the measures used and the level of analysis chosen, *Contemporary Educational Psychology*, 34(1), 77-88.

Trost, S. & Salehi-Isfahani, D. (2012). The Effect of Homework on Exam Performance: Experimental Results from Principles of Economics. *Southern Economic Journal,* 79(1), 224-242.